

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:11:45 ; Search time 22 Seconds  
(without alignments)  
4859.092 Million cell updates/sec

Title: US-10-017-216-2

Perfect score: 10490  
Sequence: 1 MLKFKYGARNPLDAGAEPI.....OLNGEIRQOVKSVLRIDYC 2053

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7491.5	71.4	1597	1 CTRO_MOUSE	P49025 mus musculus
2	6223.5	59.3	1286	1 CTRO_HUMAN	Q14578 homo sapien
3	862	8.2	631	1 DMK_MOUSE	P54265 mus musculus
4	839.5	8.0	639	1 DMK_HUMAN	Q09013 homo sapien
5	698	6.7	756	1 CBK1_YEAST	P53894 saccharomyc
6	691	6.6	469	1 ORB6_SCHPO	O13310 schizosacch
7	630	6.0	598	1 COT1_NEUCR	P38579 neurospora
8	590.5	5.6	1411	1 EEAL_HUMAN	Q15075 homo sapien
9	575.5	5.5	2238	1 G0A4_MOUSE	Q91vw5 mus musculus
10	561	5.3	2442	1 CEP2_HUMAN	Q9bv73 homo sapien
11	551	5.3	2230	1 G0A4_HUMAN	Q13439 homo sapien
12	540.5	5.2	2116	1 MYS2_DICDI	P08799 dictyosteli
13	540	5.1	444	1 KRAC_DICDI	P54644 dictyosteli
14	540	5.1	607	1 KALB_SCHPO	Q09898 schizosacch
15	538	5.1	1976	1 MYHA_RAT	Q9jlt0 rattus norv
16	536	5.1	1976	1 MYHA_BOVIN	Q27991 bos taurus
17	535.5	5.1	1433	1 REST_CHICK	Q42184 gallus gall
18	534	5.1	1790	1 USO1_YEAST	P25386 saccharomyc
19	534	5.1	1976	1 MYHB_HUMAN	P35580 homo sapien
20	533.5	5.1	3259	1 G0B1_HUMAN	Q14789 homo sapien
21	532	5.1	1407	1 TRH1_RABIT	P37709 oryctolagus
22	531.5	5.1	1875	1 MLPI_YEAST	Q02455 saccharomyc
23	529.5	5.0	2349	1 TPR_HUMAN	P12270 homo sapien
24	529	5.0	1966	1 MYSB_CABEL	P02566 caenorhabdi
25	528.5	5.0	2663	1 CENE_HUMAN	Q02224 homo sapien
26	526.5	5.0	1690	1 C190_DROME	Q9vies drosophila
27	526.5	5.0	2017	1 MYSN_DROME	Q93323 drosophila
28	524.5	5.0	584	1 DBFE_YEAST	P32328 saccharomyc
29	524.5	5.0	4684	1 PLE1_HUMAN	Q15149 homo sapien
30	524	5.0	1960	1 MYH9_HUMAN	P35579 homo sapien
31	523.5	5.0	1969	1 MYSB_CABEL	P12844 caenorhabdi
32	521	5.0	1972	1 MYHB_HUMAN	P35749 homo sapien
33	520.5	5.0	1972	1 MYHB_RABIT	P35748 oryctolagus

#### ALIGNMENTS

##### RESULT 1

CTRO\_MOUSE STANDARD; PRT; 1597 AA.  
AC P49025; 5.0 4473 1 PLE1\_CRIGR  
DT 01-FEB-1996 (Rel. 33, Created) 34 520.5 5.0 1939 1 MYH4\_HUMAN  
DT 01-FEB-1996 (Rel. 33, Last sequence update) 35 520 5.0 1939 1 MYH4\_HUMAN  
DT 28-FEB-2003 (Rel. 41, Last annotation update) 36 519.5 5.0 1938 1 MYH3\_HUMAN  
DE Citron protein (Rho-interacting, serine/threonine kinase 21). 37 517.5 4.9 1938 1 MYS\_AEQIR  
GN CIT. 38 517 4.9 1978 1 MYHB\_CHICK  
OS Mus musculus (Mouse). 39 516.5 4.9 1938 1 MYH6\_MOUSE  
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; 40 515 4.9 572 1 DBF2\_YEAST  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 41 514 4.9 1935 1 MYSS\_CYPECA  
CC NCBI\_TaxID=10090; 42 514 4.9 4687 1 PLE1\_RAT  
RN [1] 43 513.5 4.9 1939 1 MYH6\_HUMAN  
RP SEQUENCE FROM N.A. 44 513 4.9 893 1 KN8R\_YEAST  
RC TISSUE=Brain; 45 512.5 4.9 1939 1 MYH6\_MESAU  
RX MEDLINE=96128238; PubMed=8543060;  
RA Madanle P., Furuyaashiki T., Reid T., Ishizaki T., Watanabe G.,  
RA Morii N., Narumiya S.;  
RT "A novel partner for the GTP-bound forms of rho and rac.";  
RL FEBS-Lett 377:243-248(1995).  
CC FUNCTION: Putative RHO/RAC effector that binds to the GTP-bound  
CC forms of RHO and RAC1. It probably binds p21 with a tighter  
CC specificity in vivo.  
CC SUBUNIT: Homodimer (Probable).  
CC TISSUE SPECIFICITY: A major signal was observed in testis and  
CC brain, but it was also detected in thymus, spleen, kidney, heart  
CC and lung.  
CC SIMILARITY: Contains 1 CNH domain.  
CC SIMILARITY: Contains 1 PH domain.  
CC SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG  
CC binding domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U39904; AAC52341.1; -.  
PIR; S68420; S68420.  
DR MGD; MGI:105313; Cit.  
DR InterPro; IPR001180; Citron.  
DR InterPro; IPR002219; DAG\_PE-bind.  
DR InterPro; IPR001849; PH.  
DR InterPro; IPR000861; REM\_repeat.  
DR Pfam; PF00780; CNH; 1.  
DR Pfam; PF00130; DAG\_PE-bind; 1.  
DR Pfam; PF00169; PH; 1.  
DR SMART; SM00109; Cl; 1.  
DR SMART; SM00036; CNH; 1.  
DR SMART; SM00233; PH; 1.  
DR PROSITE; PS00479; DAG\_PE\_BIND\_DOM\_1; 1.  
DR PROSITE; PS00081; DAG\_PE\_BIND\_DOM\_2; 1.

DR PROSITE; PS0003; PH DOMAIN; 1.  
 KW Coiled coil; Phorbol-ester binding; SH3-binding.  
 FT DOMAIN 1 845 COILED COIL (POTENTIAL).  
 FT DOMAIN 674 870 RHO/RAC BINDING.  
 FT DOMAIN 818 821 POLY-LYS.  
 FT DOMAIN 931 979 PHORBOL-ESTER AND DAG BINDING.  
 FT DOMAIN 1011 1131 PH.  
 FT DOMAIN 1160 1457 CNH.  
 FT SITE 1521 1526 SH3-BINDING (POTENTIAL).  
 SQ SEQUENCE 1597 AA; 183448 MW; 787286C2305676DA CRC64;

Query Match 71.4%; Score 7491.5; DB 1; Length 1597;  
 Best Local Similarity 93.1%; Pred. No. 8.8e-224;  
 Matches 1492; Conservative 16; Mismatches 12; Indels 83; Gaps 4;

QY 468 MEQEMTLHRRVSEVAVLSQKEVELKASQVVEEMRLMNNQLEEDLVSAARRSDLYES 527  
 DB 9 MEQEMTLHRRVSEVAVLSQKEVELKASQVVEEMRLMNNQLEEDLVSAARRSDLYES 68  
 QY 528 VSQEDDKALQLLHDIREQSRKLOEIKQEYQAOVEEMRLMNNQLEEDLVSAARRSDLYES 587  
 DB 69 VSQEDDKALQLLHDIREQSRKLOEIKQEYQAOVEEMRLMNNQLEEDLVSAARRSDLYES 128  
 QY 588 ELRESRLAAEFKFKATECHQKLLKAKDQKQPEVGEYAKLEKINACQQLKIQLQKLEK 647  
 DB 129 ELRESRLAAEFKFKATECHQKLLKAKDQKQPEVGEYAKLEKINACQQLKIQLQKLEK 188  
 QY 648 A-----AKRAERLEKLNRRDSSEGIRKLVAEERRHSLNKKVL 691  
 DB 189 AVKASTFATELLQNIQAKAPARELEKLNRRDSSEGIRKLVAEERRHSLNKKVL 248  
 QY 692 ETWERENRUKDDIQTSQOIQMAQDKILELEKHKHREAQVSAQHLVHLKQEQHYEKI 751  
 DB 249 ETWERENRUKDDIQTSQOIQMAQDKILELEKHKHREAQVSAQHLVHLKQEQHYEKI 308  
 QY 752 KVLDNQIKDLADKETLNNMQSHHEAEKGLSEQKAMINAMDSKIRSEFQRTVELS 811  
 DB 309 KVLDNQIKDLADKETLNNMQSHHEAEKGLSEQKAMINAMDSKIRSEFQRTVELS 368  
 QY 812 EANKLAANSLSFTQRNKAQEMISELROQKFYLETQAGKLEAQNKLKEQLEKISHQDH 871  
 DB 369 EANKLAANSLSFTQRNKAQEMISELROQKFYLETQAGKLEAQNKLKEQLEKISHQDH 428  
 QY 872 SDQNRLLELTRREVSLEHEEQKLEKQLTELQSLQERESQLTALQAAARAALESQLR 931  
 DB 429 SDKSRLLELTRREVSLEHEEQKLEKQLTELQSLQERESQLTALQAAARAALESQLR 488  
 QY 932 QAKTELEETTAABEEITQALTAHRDEIQRFDAIRNSCTVITDLEBQLNQLTEDNAELNN 991  
 DB 489 QAKTELEETTAABEEITQALTAHRDEIQRFDAIRNSCTVITDLEBQLNQLTEDNAELNN 548  
 QY 992 QNFYLSKQLDEASGANDIEIVQLASEVDHLRREITEREMQLTSQKQTMELAKTCTMLBEQ 1051  
 DB 549 QNFYLSKQLDEASGANDIEIVQLASEVDHLRREITEREMQLTSQKQTMELAKTCTMLBEQ 608  
 QY 1052 VMLEALNDELLEKQEWAAWRSVLGDEKSOQECRVRELQRMLDTEKQSRARADORITES 1111  
 DB 609 VLDLEALNDELLEKQEWAAWRSVLGDEKSOQECRVRELQRMLDTEKQSRARADORITES 668  
 QY 1112 RQVVELAVXEHKAEIILALQALKEQKKAESLSDKLNLEKHAMLENNARSLOQKLETE 1171  
 DB 669 RQVVELAVXEHKAEIILALQALKEQKKAESLSDKLNLEKHAMLENNARSLOQKLETE 728  
 QY 1172 RELKQRLLEBEQAKLQQMDLQKXNHIFRLTQGGQEAALDRADLAKTBSRLDYOLENIQVLY 1231  
 DB 729 RELKQRLLEBEQAKLQQMDLQKXNHIFRLTQGGQEAALDRADLAKTBSRLDYOLENIQVLY 788  
 QY 1232 SHEKVRMEGTISQOTKLIIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKAL 1291  
 DB 789 SHEKVRMEGTISQOTKLIIDFLQAKMDQPAKKK-----VPLQYNELKAL 833  
 QY 1292 EKEKARCACAELEALQKTRIELSAREEAHRAKATDHPHPSTPATARQOIANSAIVRSPEH 1351

DB 834 EKEKARCACAELEALQKTRIELSAREEAHRAKATDHPHPSTPATARQOIANSAIVRSPEH 893  
 QY 1352 QPSAMSLIAPPSSRRKESSTPEEFSRRLKERMHNIPIHRFNVGLNMRATKCAVCLDTVHF 1411  
 DB 894 QPSAMSLIAPPSSRRKESSTPEEFSRRLKERMHNIPIHRFNVGLNMRATKCAVCLDTVHF 953  
 QY 1412 GRQASKCLECQVMCHPKCSTCLPATCGLPABYATHFTEAFCDKQNSPGLQTKEPSSSLH 1471  
 DB 954 GRQASKCLECQVMCHPKCSTCLPATCGLPABYATHFTEAFCDKQNSPGLQTKEPSSSLH 1013  
 QY 1472 LEGMMKVPNNKRGQGGWDRKYIIVLEGSKVLIDNEAREAGQREVEFECLCLPDGDSIH 1531  
 DB 1014 LEGMMKVPNNKRGQGGWDRKYIIVLEGSKVLIDNEAREAGQREVEFECLCLPDGDSIH 1073  
 QY 1532 GAVGASELANTAKA----- 1545  
 DB 1074 GAVGASELANTAKADVPYILKOWESHPTTCWPGRTLYLLAPSPDKORWVTALSVVAGG 1133  
 QY 1546 ---EKAPADAKLLGNSLLKLEGDRDLMNCTLPFSQVVLVGTTEGLYALNVLKNSLTH 1601  
 DB 1134 RVSREKAPADAKLLGNSLLKLEGDRDLMNCTLPFSQVVLVGTTEGLYALNVLKNSLTH 1193  
 QY 1602 VPGIGAVFOIYIIKDLKLMIAGERALCLVDVVKYKQSLAQSHLPAQDPDISPNIIEAV 1661  
 DB 1194 IPGIGAVFOIYIIKDLKLMIAGERALCLVDVVKYKQSLAQSHLPAQDPDISPNIIEAV 1253  
 QY 1662 KGCHLFGAGKIENGICICAAMPKSVILRYNENLSKYCIKKEITETSPCSIHTNYSIL 1721  
 DB 1254 KGCHLFGAGKIENGICICAAMPKSVILRYNENLSKYCIKKEITETSPCSIHTNYSIL 1313  
 QY 1722 IGTNKFVEIDMKQYTLSEFLDKNDHSLAPAVFASSNSFPVSIQVNSAQOREEYLLCFH 1781  
 DB 1314 IGTNKFVEIDMKQYTLSEFLDKNDHSLAPAVFASSNSFPVSIQVNSAQOREEYLLCFH 1373  
 QY 1782 EFGVFDVSYGRRSTDDLKWSRLPLAFAYREPVLFTVTHNSLEVIEIQARSSAGTAPARAY 1841  
 DB 1374 EFGVFDVSYGRRSTDDLKWSRLPLAFAYREPVLFTVTHNSLEVIEIQARSSAGTAPARAY 1433  
 QY 1842 LDIPNRYLQPAISSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPGSTSRSSPNKRG 1901  
 DB 1434 LEIPNRYLQPAISSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPGSTSRSSPNKRG 1493  
 QY 1902 PPTNEHITKRVASSPAPPGSPHPREPSTPHRY--REGTEILRRKSGPRLEREKSPG 1959  
 DB 1494 PPTNEHITKRVASSPAPPGSPHPREPSTPHRY--REGTEILRRKSGPRLEREKSPG 1553  
 QY 1960 RMLSTRERSPGRLFEDSSRGRPLPAGAVRTPLSQVNVKGRQSA 2002  
 DB 1554 RMLSTRERSPGRLFEDSSRGRPLPAGAVRTPLSQVNVKGRQSA 1596

RESULT 2  
 CTRO HUMAN  
 ID CTRO HUMAN STANDARD; PRT; 1286 AA.  
 AC O14578; Q9UPZ7;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Citron protein (Rho-interacting, serine/threonine kinase 21)  
 DE (Fragment).  
 DE CIT OR STK21 OR KIAA0949.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Connell M., Goela D., Harper M.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RC SEQUENCE OF 347-1286 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99246063; PubMed=10231032;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hiroseawa M.,  
RA Miyajima N., Tanaka A., Kozani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XIII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 6:63-70(1999).  
CC -!- FUNCTION: Putative RHO/RAC effector that binds to the GTP-bound  
CC forms of RHO and RAC1. It probably binds p21 with a tighter  
CC specificity in vivo (By similarity).  
CC -!- SIMILARITY: Contains 1 CNH domain.  
CC -!- SIMILARITY: Contains 1 PH domain.  
CC -!- SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG  
CC binding domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AC002563; AAB71327.1; --  
CC EMBL; AB023166; BAA76793.1; --  
CC Genew; HGNC:1985; CIT.  
CC MIM; 605629; --  
CC InterPro; IPR001180; Citron.  
CC InterPro; IPR002219; DAG\_PE-bind.  
CC InterPro; IPR001849; PH.  
CC InterPro; IPR000861; REM\_repeat.  
CC Pfam; PF00780; CNH; 1.  
CC Pfam; PF00130; DAG\_PE-bind; 1.  
CC Pfam; PF00169; PH; 1.  
CC SMART; SM00109; C1; 1.  
CC SMART; SM00036; CNH; 1.  
CC SMART; SM00233; PH; 1.  
CC PROSITE; PS00479; DAG\_PE\_BIND\_DOM\_1; 1.  
CC PROSITE; PS00081; DAG\_PE\_BIND\_DOM\_2; 1.  
CC PROSITE; PS00003; PH\_DOMAIN\_1  
KW Coiled coil; Phorbol-ester binding; SH3-binding.  
FT NON\_TER 1  
FT DOMAIN <1 558 COILED COIL (POTENTIAL).  
FT DOMAIN 365 561 RHO/RAC BINDING.  
FT DOMAIN 509 512 POLY-LYS.  
FT DOMAIN 622 670 PHORBOL-ESTER AND DAG BINDING.  
FT DOMAIN 702 822 PH.  
FT DOMAIN 851 1148 CNH.  
FT SITE 1212 1217 SH3-BINDING (POTENTIAL).  
FT SEQUENCE 1286 AA; 146506 MW; 498101F79EA75E85 CRC64;  
Query Match 59.3%; Score 6223.5; DB 1; Length 1286;  
Best Local Similarity 94.7%; Pred. No. 6.8e-185;  
Matches 1231; Conservative 1; Mismatches 3; Indels 65; Gaps 2;  
QY 753 VLDNQIKKOLADKETLENMQRHEEAEHKGKILSQKAMINAMDSKIRSLQRIVELSE 812  
DB 1 VLDNQIKKOLADKETLENMQRHEEAEHKGKILSQKAMINAMDSKIRSLQRIVELSE 60  
QY 813 ANKLAANSSIFTORNNKAQEMISELRQOKFYLETQAGKLEAQNKLBEQLEKISHQDHS 872  
DB 61 ANKLAANSSIFTORNNKAQEMISELRQOKFYLETQAGKLEAQNKLBEQLEKISHQDHS 120  
QY 873 DKNRILLETRLEVSLEHEEQKLEKRLQTELQLSQRESQLTALQAARALESQLRQ 932  
DB 121 DKNRILLETRLEVSLEHEEQKLEKRLQTELQLSQRESQLTALQAARALESQLRQ 180  
QY 933 AKTELETTAAEETIQAHTAHDEIQRKFDALRNSCTVITDLEEQNLQITDNEALNNQ 992  
DB 181 AKTELETTAAEETIQAHTAHDEIQRKFDALRNSCTVITDLEEQNLQITDNEALNNQ 240  
QY 993 NFVLSKQDDEASGANDEIVQLRSEVDHLRREITERENQOLTSQKTMEALKTCTMLEEQV 1052  
DB 241 NFVLSKQDDEASGANDEIVQLRSEVDHLRREITERENQOLTSQKTMEALKTCTMLEEQV 300

RESULT 3

DMK\_MOUSE

ID DMK\_MOUSE STANDARD; PRT; 631 AA.

QY 1053 MDLEALNDELLEKERQWEAWRSVLGDEKSFQECRVRELOQMLDTEKQSRARADQRTESR 1112  
DB 301 MDLEALNDELLEKERQWEAWRSVLGDEKSFQECRVRELOQMLDTEKQSRARADQRTESR 360  
QY 1113 QVVELAVKEHKAETILALQQALKQKKAESLSKLANDLEKHAMLENMARSLOOKLETER 1172  
DB 361 QVVELAVKEHKAETILALQQALKQKKAESLSKLANDLEKHAMLENMARSLOOKLETER 420  
QY 1173 ELKQRLLEEQAQLQQQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYOLENIQVLYS 1232  
DB 421 ELKQRLLEEQAQLQQQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYOLENIQVLYS 480  
QY 1233 HEKVMGTTISQOTKLIDFLQAKMDQPAKKKGLFSRKKEDPALPTQVPLQYNELKLALE 1292  
DB 481 HEKVMGTTISQOTKLIDFLQAKMDQPAKKKGLFSRKKEDPALPTQVPLQYNELKLALE 525  
QY 1293 KEKARCAEAEALQKTRIELRSAREEAAHRAKADHPHPSTPATARQOIAMSAIVRSPEHQ 1352  
DB 526 KEKARCAEAEALQKTRIELRSAREEAAHRAKADHPHPSTPATARQOIAMSAIVRSPEHQ 585  
QY 1353 PSAMSLIAPPSSRRKESSTPEFSRRLKERMHNIHPRFNVGLNMRAKCAVCLDTHVFG 1412  
DB 586 PSAMSLIAPPSSRRKESSTPEFSRRLKERMHNIHPRFNVGLNMRAKCAVCLDTHVFG 645  
QY 1413 ROASKCLEQVMCHPKCSTCLPATCGLPABYATHFTFAFCRDKNMSPGLQTKPSSSLHL 1472  
DB 646 ROASKCLEQVMCHPKCSTCLPATCGLPABYATHFTFAFCRDKNMSPGLQTKPSSSLHL 705  
QY 1473 EGMKVPNNKRGQGWDRKYIVLEGSKVIYDNEAREAGQRPVEEFELCLPDGDSVING 1532  
DB 706 EGMKVPNNKRGQGWDRKYIVLEGSKVIYDNEAREAGQRPVEEFELCLPDGDSVING 765  
QY 1533 AVGASELAN TAKA----- 1545  
DB 766 AVGASELAN TAKADVPIVKMESHPTTQWPGRTLYLLAPFPDKQWVTALSVVAGGR 825  
QY 1546 ---EKAADAKLLGNSLLKLEGGDRLDNMCTLPFDQVLVGTVEEGLYALNVLKNSLTHV 1602  
DB 826 VSREKAADAKLLGNSLLKLEGGDRLDNMCTLPFDQVLVGTVEEGLYALNVLKNSLTHV 885  
QY 1603 PGIGAVFQYIIKOLEKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIFFAVK 1662  
DB 886 PGIGAVFQYIIKOLEKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIFFAVK 945  
QY 1663 GCHLFGAGKTEGLCI CAAMPSKVILRYNENLSKYCIKKEIETSEPCSCIHFNTYSILI 1722  
DB 946 GCHLFGAGKTEGLCI CAAMPSKVILRYNENLSKYCIKKEIETSEPCSCIHFNTYSILI 1005  
QY 1723 GTNKFYEDMKQYTLLEBFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGOREEYLLCFHE 1782  
DB 1006 GTNKFYEDMKQYTLLEBFLDKNDHSLAPAVFAASSNSFPVSIVQVNSAGOREEYLLCFHE 1065  
QY 1783 FGVFVDSVGRSRTDDLKWSRLPLAFAYREPFLVTHFNSLVEIEIQARSAGTAPAYL 1842  
DB 1066 FGVFVDSVGRSRTDDLKWSRLPLAFAYREPFLVTHFNSLVEIEIQARSAGTAPAYL 1125  
QY 1843 DIPNPRYLGPATSSGAIY LASSYQDKLRVICCKGNLVKESGTEHHRGPGSTSRSSPNKRG 1902  
DB 1126 DIPNPRYLGPATSSGAIY LASSYQDKLRVICCKGNLVKESGTEHHRGPGSTSRSSPNKRG 1185  
QY 1903 PTYNEHTKRVASSPAPPEGSHPREPSTPHRYEGTELRDCKSPGRPLEREKSPGRML 1962  
DB 1186 PTYNEHTKRVASSPAPPEGSHPREPSTPHRYEGTELRDCKSPGRPLEREKSPGRML 1245  
QY 1963 STRRERSGRLFEDESSRGLPAGAVRTPLSQVKNKGQSA 2002  
DB 1246 STRRERSGRLFEDESSRGLPAGAVRTPLSQVKNKVDQSS 1285

```

P54265;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Myotonic-protein kinase (EC 2.7.1.-) (Myotonic dystrophy protein
kinase) (MDPK) (DM-kinase) (DMK) (DMPK) (MT-PK).
DR DMPK OR MDPK OR DM15.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=93271990; PubMed=8499920;
RA Mahadevan M.S., Amemiya C., Jansen G., Sabourin L., Baird S.,
RA Neville C.E., Wormskamp N., Segers B., Batzer M., Lamerdin J.,
RA de Jong P.J., Wieringa B., Kornelijs R.G.;
RT "Structure and genomic sequence of the myotonic dystrophy (DM kinase)
RT gene.";
RL Hum. Mol. Genet. 2:299-304(1993).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=129/SV;
RX MEDLINE=94004865; PubMed=8401505;
RA Jansen G., Bartolonei M., Kalscheuer V., Merckx G., Wormskamp N.,
RA Mariman E., Smeets D., Ropers H.-H., Wieringa B.;
RT "No imprinting involved in the expression of DM-kinase mRNAs in mouse
RT and human tissues.";
RL Hum. Mol. Genet. 2:1221-1227(1993).
RN [3]
SEQUENCE OF 117-631 FROM N.A. (ISOFORMS 2; 4; 5 AND 8).
RP TISSUE=Brain;
RX MEDLINE=93251003; PubMed=1302022;
RA Jansen G., Mahadevan M.S., Amemiya C., Wormskamp N., Segers B.,
RA Hendriks W., O'Hoy K., Baird S., Sabourin L., Lennon G., Jap P.L.,
RA Iles D., Coerwinkel M., Hofker M., Carrano A.V., de Jong P.J.,
RA Kornelijs R.G., Wieringa B.;
RT "Characterization of the myotonic dystrophy region predicts multiple
RT protein isoform-encoding mRNAs.";
RL Nat. Genet. 1:261-266(1992).
RN [4]
ALTERNATIVE SPLICING (ISOFORMS 1; 3; 5; 6; 8 AND 9).
RX MEDLINE=20164299; PubMed=10699184;
RA Groenen P.J.T.A., Wansink D.G., Coerwinkel M., van den Broek W.,
RA Jansen G., Wieringa B.;
RT "Constitutive and regulated modes of splicing produce six major
RT myotonic dystrophy protein kinase (DMPK) isoforms with distinct
RT properties.";
RL Hum. Mol. Genet. 9:605-616(2000).
CC -!- FUNCTION: May play a role in intracellular communication.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=10;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P54265-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P54265-2; Sequence=VSP_004819;
CC Name=3;
CC IsoId=P54265-3; Sequence=VSP_004820;
CC Name=4;
CC IsoId=P54265-4; Sequence=VSP_004820, VSP_004821, VSP_004822;
CC Name=5;
CC IsoId=P54265-5; Sequence=VSP_004823, VSP_004824;
CC Name=6;
CC IsoId=P54265-6; Sequence=VSP_004820, VSP_004823, VSP_004824;
CC Name=7;
CC IsoId=P54265-7; Sequence=VSP_004825, VSP_004826;
CC Name=8;
CC IsoId=P54265-8; Sequence=VSP_004827;
CC Name=9;
CC IsoId=P54265-9; Sequence=VSP_004820, VSP_004827;
CC Name=10;
CC IsoId=P54265-10; Sequence=VSP_004828, VSP_004829;

```

---

```

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINATED TO DATE,
CC WITH A PREDOMINANCE IN BRAIN, SKELETAL MUSCLE, HEART, AND OTHER
CC TISSUES CONTAINING SMOOTH MUSCLE.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z38015; CRA86113.1; -
CC EMBL; Z21503; CRA79715.1; -
CC EMBL; Z21504; CRA79715.1; JOINED.
CC EMBL; Z21505; CRA79715.1; JOINED.
CC EMBL; Z21506; CRA79715.1; JOINED.
CC EMBL; S60313; AAC60667.1; -
CC EMBL; S60314; AAC60666.1; -
CC EMBL; S60315; AAC60665.1; -
CC EMBL; S60316; AAC60664.1; -
CC PIR; S71829; S71829.
CC HSP; P05132; 15KX.
CC MGD; MGI:94906; Dm15.
CC InterPro; IPR000961; Pkinase_C.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00133; S_TK_X; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
CC Coiled coil; Alternative splicing.
CC DOMAIN 71 339 PROTEIN_KINASE.
CC NP_BIND 77 85 ATP (BY SIMILARITY).
CC BINDING 100 100 ATP (BY SIMILARITY).
CC DOMAIN 464 532 COILED COIL (POTENTIAL).
CC ACT_SITE 195 195 BY SIMILARITY.
CC DOMAIN 597 604 POLY-ALA.
CC VARSPLIC 328 356 Missing (in isoform 2).
CC FT FTId=VSP_004819.
CC FT Missing (in isoform 3, isoform 4, isoform
CC FT 6 and isoform 9).
CC FT FTId=VSP_004820.
CC FT DNOVP -> LKRPT (in isoform 4).
CC FT FTId=VSP_004821.
CC FT Missing (in isoform 4).
CC FT FTId=VSP_004822.
CC FT AI -> DP (in isoform 5 and isoform 6).
CC FT FTId=VSP_004823.
CC FT Missing (in isoform 5 and isoform 6).
CC FT FTId=VSP_004824.
CC FT AITGVSPRATDPHSLDGPAA -> GESLTCFQPRGHWE
CC FT MGGMLGV (in isoform 7).
CC FT FTId=VSP_004825.
CC FT Missing (in isoform 7).
CC FT FTId=VSP_004826.
CC FT VARSPLIC 552 631
CC FT LDGPPAVAVGQ -> ASRQLPKGTP (in isoform
CC FT 10).
CC FT FTId=VSP_004828.
CC FT Missing (in isoform 10).
CC FT FTId=VSP_004829.

```

```

SQ SEQUENCE 631 AA; 69601 MW; 5EB800A37EA81DF4 CRC64;
Query Match 8.2%; Score 862; DB 1; Length 631;
Best local similarity 36.6%; Pred. No. 2.3e-20;
Matches 199; Conservative 104; Mismatches 191; Indels 50; Gaps 13;

QY 46 LSGREGILDALFVLFEECSQPALMKIKHVSFNVRKYSDTIAELQELQPSAKDFEVSRLVGC 105
Db 20 LGLEPFLDLLGVHQLGSHALQADKYVADFQWPEPIAARLKEVRLQDRDDFEILKVIGR 79
QY 106 GHFAEVQVVRKATGDIYAKVMKKKALLAQSQVFFEEERNILSRSTSPWLPQLOAFQ 165
Db 80 GAFSEVAVVMKQTGTVYAKIMKMKWMLKRGVSCFEERDVLVKGDRMTWLQHFAPQ 139
QY 166 DKNHLYMEEYOPGGDLILNARYEDQDENLIQFYLAELIIVASHVLMGVVHRDIKPE 225
Db 140 DENILYVMEVYVGGDLILNARYEDQDENLIQFYLAELIIVASHVLMGVVHRDIKPD 199
QY 226 NILVDRTHIKLVDFGSAAMNSNMKNVNAKLPIGTPDYMAPEVLTVMGDQKGYGLDC 284
Db 200 NILDCRCHIRLADFGSLKLPQDGMVRSVAVGTPDYLSPILQAVGGPGAGSYGPEC 259
QY 285 DWSVGVIAEYIMYGRSPAEGTSARTFNINMNFQFLKFPDDPKVSSDFLDLILQSLLC 343
Db 260 DMWALGVFAYEYFGTTPYADSTAEYAKIYHREHLSLPLADTVVPEAODLIRGLLC 319
QY 344 GOKERLKFEGL--LCCHPFFSKIDMNIIRNSPPFVPTLKSDDDTSNFD--BPEKNSMVS 398
Db 320 PAEIRLGRGAGDFQKHFFFGDLWGLRDSVPFPTDFEGATDTCNPDVVEDRLTAMVS 379
QY 399 SSPCQLSPSGFS---GEELPFVGFYSKALGILGRSESVSGLDSPAKTSSMEKILLIKS 455
Db 380 GGGETLSDMQEDMPLGVRLFFVGYSY---CCMAFRDNQV-----PDPTFMBLEALQUPV 430
QY 456 KELO--DSQDKCHKMEQ-----EMTLHRRVSEVEAVLSQKVELKASRTQSLLEQDLIA 508
Db 431 SDLQGLDLPVSPPPQVAAEDLVAVPAPVAAETTVTLQLO-----EALSEEVIL 482
QY 509 TYITECSLKSRLSEARMEVQSDKALQLLHDIREQSKRLQEIKEQYQAOVEMRLMM 568
Db 483 TR---QSLRELEAIR-TANQNFSSQLQ-----EAEVRNRDLAEHVRLQIQRMM 526
QY 569 NQLE 572
Db 527 EMLQ 530

RESULT 4
DMK_HUMAN
ID DMK_HUMAN STANDARD; PRT; 639 AA.
AC Q09013; Q16205;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myotonin-protein kinase [EC 2.7.1.-] (Myotonic dystrophy protein
DE kinase) (MDPK) (DM-kinase) (DMK) (DMPK) (MT-PK).
GN DMPK OR MDPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92188202; PubMed=1546326;
RA Fu Y.-H., Pizzuti A., Fenwick R.G. Jr., King J., Rajnarayan S.,
RA Dunne P.W., Dubel J., Nasser G.A., Ashizawa T., de Jong P.J.,
RA Wieringa B., Korneluk R., Perryman M.B., Epstein H.F., Caskey C.T.;
RT "An unstable triplet repeat in a gene related to myotonic muscular
RT dystrophy."
RL Science 255:1256-1258 (1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 5; 6; 7 AND 8).
RX MEDLINE=93227030; PubMed=8469976;

```

```

RA Fu Y.-H., Friedman D.L., Richards S., Pearlman J.A., Gibbs R.A.,
RA Pizzuti A., Ashizawa T., Perryman M.B., Scariato G., Fenwick R.G. Jr.,
RA Caskey C.T.;
RT "Decreased expression of myotonin-protein kinase messenger RNA and
RT protein in adult form of myotonic dystrophy."
RL Science 260:235-238 (1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 9; 10; 11 AND 12).
RX MEDLINE=93271990; PubMed=8499920;
RA Mahadevan M.S., Amemiya C., Jansen G., Sabourin L., Baird S.,
RA Neville C.E., Wormskamp N., Segers B., Batzer M., Lamerdin J.,
RA de Jong P.J., Wieringa B., Korneluk R.G.;
RT "Structure and genomic sequence of the myotonic dystrophy (DM kinase)
RT gene."
RL Hum. Mol. Genet. 2:299-304 (1993).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 11).
RC TISSUE=Brain, and Muscle;
RX MEDLINE=94140369; PubMed=7905855;
RA Shaw D.J., McCurrach M., Rundle S.A., Harley H.G., Crow S.R., Sohn R.,
RA Thirion J.-P., Hamshire M.G., Buckler A.J., Harper P.S., Housman D.E.,
RA Brook J.D.;
RT "Genomic organization and transcriptional units at the myotonic
RT dystrophy locus."
RL Genomics 18:673-679 (1993).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 11).
RC TISSUE=Muscle;
RX MEDLINE=94357271; PubMed=8076686;
RA Sasagawa N., Sorimachi H., Maruyama K., Arahata K., Ishiura S.,
RA Suzuki K.;
RT "Expression of a novel human myotonin protein kinase (McpK) cDNA clone
RT which encodes a protein with a thymopoleitin-like domain in COS
RT cells."
RL FEBS Lett. 351:22-26 (1994).
RN [6]
RP SEQUENCE OF 24-639 FROM N.A. (ISOFORMS 9 AND 11).
RC TISSUE=Brain, and Heart;
RX MEDLINE=93251003; PubMed=1302022;
RA Jansen G., Mahadevan M., Amemiya C., Wormskamp N., Segers B.,
RA Hendriks W., O'Hoy K., Baird S., Sabourin L., Lennon G., Jap P.L.,
RA Iles D., Coerwinkel M., Hofker M., Carrano A.V., de Jong P.J.,
RA Korneluk R.G., Wieringa B.;
RT "Characterization of the myotonic dystrophy region predicts multiple
RT protein isoform-encoding mRNAs."
RL Nat. Genet. 1:261-266 (1992).
RN [7]
RP SEQUENCE OF 52-639 FROM N.A. (ISOFORM 13).
RX MEDLINE=92154692; PubMed=1310900;
RA Brook J.D., McCurrach M., Harley H.G., Buckler A.J., Church D.,
RA Aburatani H., Hunter K., Stanton V.P., Thirion J.-P., Hudson T.,
RA Sohn R., Zemelman B., Snell R.G., Rundle S.A., Crow S., Davies J.,
RA Shelbourne P., Buxton J., Jones C., Juvenon V., Johnson K.,
RA Harper P.S., Shaw D.J., Housman D.E.;
RT "Molecular basis of myotonic dystrophy: expansion of a trinucleotide
RT (CTG) repeat at the 3' end of a transcript encoding a protein kinase
RT family member."
RL Cell 68:799-808 (1992).
RN [8]
RP SEQUENCE OF 539-575 FROM N.A. (ISOFORM 14).
RC TISSUE=Brain;
RX MEDLINE=96063635; PubMed=7488138;
RA Gennarelli M., Lucarelli M., Zelano G., Pizzuti A., Novelli G.,
RA Dallapiccola B.;
RT "Different expression of the myotonin protein kinase gene in discrete
RT areas of human brain."
RL Biochem. Biophys. Res. Commun. 216:489-494 (1995).
CC -!- FUNCTION: May play a role in intracellular communication.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=12;
CC Name=1;
CC Name=2;
CC IsoId=Q09013-1; Sequence=Displayed;
CC Name=2;

```



Db 528 RLOQRMELLOAE 540

RESULT 5

CBK1 YEAST

ID CBK1 YEAST STANDARD; PRT; 756 AA.

AC P53894; 1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serine/threonine-protein kinase CBK1 (EC 2.7.1.1.) (cell wall

DE biosynthesis kinase)

GN CBK1 OR YN1161W OR N1272.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Schizosaccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / FY1679;

RX MEDLINE=96287653; PubMed=8686380;

RA Nasr F., Becam A.-M., Herbert C.J.;

RT "Pagip, a novel protein associated with protein kinase Cbx1p, is

RT required for cell morphogenesis and proliferation in Saccharomycetes

RT cerevisiae";

RL Mol. Biol. Cell 13:503-514 (2002).

CC -1- FUNCTION: Protein kinase that seems to play a role in thr

CC regulation of cell morphogenesis and proliferation.

CC -1- SUBUNIT: Associates with PAG1/TAO3.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC COT1 SUBFAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

EMBL; X92517; CAA63278.1; -.

EMBL; Z71437; CAA96048.1; -.

PIR; S60966; S60966.

GermOnline; 143167; -.

SGD; S0005105; CBK1.

GO; GO:0005935; C:bud neck; IDA.

GO; GO:0005737; C:cytoplasm; IDA.

GO; GO:0005634; C:nucleus; IDA.

GO; GO:0000753; P:cellular morphogenesis during conjugation w. . . ; IDA.

GO; GO:0030012; P:establishment and/or maintenance of cell po. . . ; IMP.

GO; GO:0007096; P:regulation of exit from mitosis; IMP.

InterPro; IPR00961; Pkinase C.

InterPro; IPR00719; Prot\_kinase.

InterPro; IPR008271; Ser\_Thr\_pkin\_A5.

InterPro; IPR002290; Ser\_Thr\_pkinase.

Pfam; PF00069; pkinase 1.

Pfam; PF00433; pkinase C; 1.

ProDom; PD000001; Prot\_kinase; 2.

SMART; SM00133; S\_TK\_X; 1.

SMART; SMO0220; S\_TK; 1.

PROSITE; PS00107; PROTEIN KINASE ATP; 1.

PROSITE; PS00108; PROTEIN KINASE ST; 1.

PROSITE; PS00011; PROTEIN KINASE DOM; 1.

Transferase; Serine/threonine-protein kinase; ATP-binding.

FT DOMAIN 23 34 POLY-GLN.

FT DOMAIN 160 164 POLY-SER.

FT DOMAIN 214 232 POLY-GLN.

FT DOMAIN 235 250 POLY-GLN.

FT DOMAIN 352 672 PROTEIN KINASE.

FT NP BIND 358 366 ATP (BY SIMILARITY).

FT BINDING 381 381 ATP (BY SIMILARITY).

FT ACT SITE 475 475 BY SIMILARITY.

SQ SEQUENCE 756 AA; 86946 MW; 87EBDCD2C3C96EE11 CRC64;

Query Match 6.7%; Score 698; DB 1; Length 756;

Best Local Similarity 32.5%; Pred. No. 3e-15;

Matches 162; Conservative 83; Mismatches 159; Indels 94; Gaps 10;

QY 4 PKYCARNP--LDAGAAPPIASRLNLFQCKPPFMTQQMSPLSREGILDALFVLFEE 61

DB 265 YMYFERPDLLTKGTQDKAAVKLKIEFYQSSVKYAIERNRVERVELETLSHWSEER 324

QY 62 CSQPALMKIKHVSFNVRKYSDTIAELOPQSAKDFEVRSLVGGCHFAEQVYVREKATGD 121

DB 325 KSRQLSSLGKESQFLR-----LRTRLSLEDFHTVKVIGKGAFAEVRLLVQKDTGK 376

QY 122 IYAKMKKKALLAQEQVSPFEERNILSRSTSPWIPQLOVAFODKXHLVLMEEYQPGGD 181

DB 377 IYAKMTLKSEMYKKDQLAHVAERDVLAGSDSPWVSLYSPQDAQYLYLIMEFLPGGD 436

QY 182 LLSLNRYEDQDENLIQFVLAELILAVHSVHLMGYVHRDIKPENILVDRTHGHIKLVDFG 241

DB 437 LMTMLIRWQ-LFTEDVTRFYMAECILAIETHIKLGFHRIKDPNINILDIRGHIKLSDFG 495

QY 242 -----SAKXNSN----- 249

DB 496 LSTGFHKTHDSNYKKLLQDQENATNG:SKPGTYNANTDTANKRQTWVVDISISITMSNRQ 555

QY 250 -----KMNNAKLPITGPDYMAPEVLTVMNGDGKGTGYGLDCDWSVGVYIAYEMYGRS 301

DB 556 QIQWRKSRRLMAYSTVGPDIAPFIYOG-----YQECDDWSLGAIMYECLIGWP 609

QY 302 PFAEGTSARTFNINMPORFLKFPDDPKVSSDFLDLIQSLCQKQERLKFEF-----LCCH 357

DB 610 PFCSETPOETRYKMNFEQTLQFPDDIHISYAEELIRLLTHADQRLGRHGGADEIKSH 669

QY 358 PFESKIDWNINRNPPPPVPVPLTKSDDTSNFDPEKSNWSVSSPC-----QLSPSG 408

DB 670 PFFRCVDWNTIRQVEAPYIPKLSITDTRFFPTDELEN-VFDSPAMAQAQKQREMTKQG 728

QY 409 FSG-----EELPFVGFYSYK 423

DB 729 GSAPVKEDLPFGITYSR 746

RESULT 6

ORB6 SCHPO

ID ORB6 SCHPO STANDARD; PRT; 469 AA.

AC O13310;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serine/threonine-protein kinase orb6 (EC 2.7.1.37).

GN ORB6 OR SPAC821.12.

OS Schizosaccharomycetes pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98301608; PubMed=9636183;

RA Verde F., Wiley D.J., Nurse P.;

RT "Fission yeast orb6, a ser/thr protein kinase related to mammalian

RT rho kinase and myotonic dystrophy kinase, is required for maintenance

RT of cell polarity and coordinates cell morphogenesis with the cell

RT cycle.";





FT endosome antigen 1.";

RL J. Biol. Chem. 278:8494-8500(2003).

RN [10]

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1289-1411 IN COMPLEX WITH

RP PHOSPHATIDYLINOSITOL 3-PHOSPHATE, AND HOMODIMERIZATION.

RX MEDLINE=21617582; PubMed=11741531;

RA Dumas J.J., Merithew E., Sudharshan E., Rajamani D., Hayes S.,

RA Lawe D., Corvera S., Lambright D.G.;

RT "Multivalent endosome targeting by homodimeric BEA1.";

RL Mol. Cell 8:947-958(2001).

RN [11]

RP STRUCTURE BY NMR OF 1346-1410 ALONE AND IN COMPLEX WITH

RP PHOSPHATIDYLINOSITOL 3-PHOSPHATE.

RX MEDLINE=21143489; PubMed=11230696;

RA Kutateladze T.G., Overduin M.;

RT "Structural mechanism of endosome docking by the FYVE domain.";

RL Science 291:1793-1796(2001).

CC -!- FUNCTION: Binds phospholipid vesicles containing

CC phosphatidylinositol 3-phosphate and participates in endosomal

CC trafficking.

CC -!- SUBUNIT: Homodimer. Binds STX6. Binds RAB5A, RAB5B, RAB5C and

CC RAB22A that have been activated by GTP-binding.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic and peripheral membrane protein

CC associated with early endosomes.

CC -!- DOMAIN: The FYVE-type zinc finger domain mediates interactions

CC with phosphatidylinositol 3-phosphate.

CC -!- DISEASE: Antibodies against BEA1 are found in sera from patients

CC with subacute cutaneous lupus erythematosus and other autoimmune

CC diseases.

CC -!- SIMILARITY: Contains 1 C2H2-type zinc finger.

CC -!- SIMILARITY: Contains 1 FYVE-type zinc finger.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; L40157; AAA79121.1; -.

DR EMBL; X78998; CAA55632.1; -.

DR PIR; A57013; A57013.

DR PDB; 1HY1; 14-MAR-01.

DR PDB; 1HYJ; 14-MAR-01.

DR PDB; 1JOC; 28-DEC-01.

DR Genew; HGNC:3185; BEA1.

DR MIM; 605070; -.

DR GO; GO:0005737; C:cytoplasm; TAS.

DR GO; GO:0005769; C:early endosome; TAS.

DR GO; GO:0005624; C:membrane fraction; TAS.

DR GO; GO:0005345; F:phosphatidylinositol binding; TAS.

DR GO; GO:0006899; P:nonselective vesicle transport; TAS.

DR InterPro; IPR007087; ZnF\_C2H2.

DR InterPro; IPR003036; ZnF\_FYVE.

DR Pfam; PF01363; FYVE; 1.

DR Pfam; PF00096; zf-C2H2; 1.

DR SMART; SM00064; FYVE; 1.

DR SMART; SM00355; ZnF\_C2H2; 1.

DR PROSITE; PS01078; ZF\_FYVE; 1.

DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.

DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 1.

KW Antigen; Metal-binding; Zinc; Coiled coil; Zinc-finger; 3D-structure.

FT DOMAIN 74 1348

FT COILED COIL (POTENTIAL).

FT GLN/GLU/LYS-RICH.

FT GLN/GLU/LYS-RICH.

FT GLU/LYS-RICH.

FT C2H2-TYPE.

FT FYVE-TYPE.

FT E-3A: STRONGLY REDUCES INTERACTION WITH

FT RAB5C.

FT F-3A: STRONGLY REDUCES INTERACTION WITH

FT RAB5C.

FT MUTAGEN 42 42 I->A: STRONGLY REDUCES INTERACTION WITH

FT RAB5C.

FT P->A: STRONGLY REDUCES INTERACTION WITH

FT RAB5C.

FT M->A: STRONGLY REDUCES INTERACTION WITH

FT RAB5C.

FT Y->A: STRONGLY REDUCES INTERACTION WITH

FT RAB5C.

FT W->A: REDUCES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT D->V: REDUCES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT N->D: REDUCES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT C->S: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT F->A: STRONGLY REDUCES

FT PHOSPHATIDYLINOSITOL 3-PHOSPHATE BINDING

FT AND ENDOSOMAL LOCATION.

FT VT->EE,GG: ABOLISHES

FT PHOSPHATIDYLINOSITOL 3-PHOSPHATE BINDING

FT AND ENDOSOMAL LOCATION.

FT R->A: ABOLISHES ENDOSOMAL LOCATION.

FT R->A: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT H->A: ABOLISHES ENDOSOMAL LOCATION.

FT H->A: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT C->A: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT R->G: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT C->A: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT G->A: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT C->A: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT R->G: STRONGLY REDUCES

FT PHOSPHATIDYLINOSITOL 3-PHOSPHATE BINDING

FT AND ABOLISHES ENDOSOMAL LOCATION.

FT C->S: ABOLISHES PHOSPHATIDYLINOSITOL 3-

FT PHOSPHATE BINDING AND ENDOSOMAL LOCATION.

FT C->S (IN REF. 1).

FT LQ ->S (IN REF. 1).

FT A ->S (IN REF. 1).

FT A ->R (IN REF. 1).

FT D ->E (IN REF. 1).

FT EQ ->DE (IN REF. 1).

FT

FT MUTAGEN 1400 1400

FT MUTAGEN 1405 1405

FT CONFLICT 255 255

FT CONFLICT 258 259

FT CONFLICT 277 277

FT CONFLICT 284 284

FT CONFLICT 520 520

FT CONFLICT 575 576

FT

Query Match 5.6%; Score 590.5; DB 1; Length 1411;

Best Local Similarity 22.4%; Pred. No. 1.le-11;

Matches 269; Conservative 238; Mismatches 410; Indels 285; Gaps 47;

QY 455 SKELQDSQDKCHKMEQEMTLHRRVSEVAVLSQKVELKASQTQ-----RSLLEQDL 507

Db 106 SEELKKELEKYQGLQOQEAQKPGDLVTSSAEQLSQLEQLEEAQTFNPNKQMDLPQKA 165

QY 508 ATYITECS-----SLKRSLEQARMEVSQEDDKALQLLHDIR----- 543

Db 166 AQLATEIADTKSKYDEERSLEAEAKVTRTLBELNKEATVIQDLKLTTELQRFIEDVAV 225

QY 544 -----EQSRKLQEIKEQ--EYQAOQVEEMELMNQLEEDLVARRRSDL 584

Db 226 LKXELVQVQTLMDNMWTLERERESEKLDCKEKLQSQVASEATISQLRSELAKPGQEVAV 285

QY 585 YESELRESRLAAEFKPKKATECQHKLLKAKDQKQPEVGEYAKLEKINAQ----- 634

Db 286 YVQELQKLKSVNELTQKNTLTENLLK-KEQ-----DYTKLEEKHEESVSKQIQAT 338

QY 635 ---QLKXIQLEQLEKAAKERAERELEKLQNRDSEGIKRLKLEAEERHLENKVKR 690

Db 339 LHQKDLCCQQLQSLR-SASETSLSLRHVELSEKGEATQKLKEELSEVETKYQLKAEFKQ 397

```
QY 691 LETMERENRLKDDIOTKSOIQOMADKILELE---EKHREAQVSAQHLEVLHKKQEQH 746
DQ 398 LQ---QCEEEKEHGLQLOS-EINQLHSLKLETERQLGEAHLKQKQKQRLSSKLMDKQKQ 454
QY 747 YEE---KIVULNQIKKDLADKETLENMQRHBEBAHEKGKILSKOKAMINAMDSKIR-- 801
DQ 455 VADLQKLSRLBEQLKEKVTNSTELOHLDKTKQK-HQ-----EQOALQOQTTAKLREA 507
QY 802 --SLEQRIVELGEANKLAANSLSFTORNKKAQEMISELRQOK--FYLETOAG----- 850
DQ 508 QNDLEQVLRQIGDKQKIQNLKALLOKS-----KENISLLEKREBDLYAKIQAGEGETAVL 563
QY 851 -KLEAQRNRLBQL-----EKISHQDSK-----NLLLELETRL 884
DQ 564 NQLQKXNHTLQEQVTQLTEKLNQKSHKQAOENLHDQVQEQKAHLRAAQDVLSTLSV 623
QY 885 REVSLSEHEQK-----LELXQLELOLSLQ-----ERESOLITALQARAALSESQ 929
DQ 624 NELNSQINESKVKVSOLDIQIKAK-TELLLSBAKTAQADLQNLHDLTAQNALQKQKE 682
QY 930 LRQAKTELEBETTA---EABEEIQAHTADEIQRFDALENSCTVITDLEEQNLQLTED- 985
DQ 683 LNKITTQLDQVTAQLQDKQKSHQSLKESHLKKEYKVSLEQK---TEELEGQIKLEADS 739
QY 986 -----NAILNNQNFYLSKOLDEASGANDIV-----QLRSE-V 1017
DQ 740 LEVKASKEQALQDLOQQRQNTDLELRATLSKQLE---MEKEIVSSVTRLDLQKKSSEAL 795
QY 1018 DHLRREITEREMQLTQSQTMBEAL-----KTTCTMLEEQVMDLEALNDEL 1062
DQ 796 ESKQKLTQEBEKQKQLKQDFETLSQETKIQHELNRRICVTIVTELQKVKMEKALMTL 855
QY 1063 LEKERQWEAARNSVLGDEKQCFQCRVRELQK---MLDTEK-----QSRARADQRIYES 1111
DQ 856 STVKDKLSKVSQSLKNSKSEFE---KENQKRAAILDLEKTKELKHQLOVQOMENTLREQ 912
QY 1112 RQVVELAVKEHKA-----EILALQALKEQKLEKESLSDKLNDEKXHAMLE 1158
DQ 913 KELKKSLEKEKASHQKLELNSMQBOLQAOQNTLKQNEKESQOQGNINEUKQSSEBQK 972
QY 1159 MNARSLO-----QKLETERELKQRLLEBAKLOQQMD-----LQKNH-----FRLTQ 1201
DQ 973 KQIEALQELGKIAVLQKTELENKLOQLTOAQELAAEKEKISVLQNNVKEQSEFQKLO 1032
QY 1202 GLQEALDRADLLKTERSDELYQENIQVLYSHKVKMEGTISQOQKLDLFL---QAKMDQ 1258
DQ 1033 --SDFYGRESELLATQDLKSVEEKLS-LAQEDLISNRNQIGNQNKLIQELKTAKATLEQ 1089
QY 1259 -PAKKKGLFSRRKEDPALPTQVPLQYNELKLALKEKAKARCAELBEALQKTRIELRSARE 1317
DQ 1090 DSAKKEQQLQERCKALQDQKESLKEKE---LVNEKSKLAEIIBEIKRQKEKITKLE 1145
QY 1318 EAAHRKATDHPHPSTPATRQOIAMSAIVRSEPHQPSAMSLIAPSS-----R 1365
DQ 1146 ELKSHLESIKETITNLDKAKQLL---IQCKLELOKADSLKAAVEQEKRNOQIILKQVYK 1201
QY 1366 RKESSIPESFRRRLKERMHNHPIHRENVG-----LNMRAKCAVCLDTVHFGRQA 1415
DQ 1202 KESEELKKEFIEK-EAKLHSEIKKEK-EVGNKXHEENEAKLTQITNALNENLGTIVKKEWQS 1259
QY 1416 SK 1417
DQ 1260 SQ 1261
```

## RESULT 9

```
GOA4_MOUSE STANDARD; PRT; 2238 AA.
AC Q91VW5; O70365; Q8CGH6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
```

```
DE GN Golgi autoantigen, golgin subfamily A member 4 (tGolgin-1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=22153218; PubMed=12162805;
RA Cowan D.A., Gay D., Bieler B.M., Zhao H., Yoshino A., Davis J.G.,
RA Tonayko M.M., Murali R., Greene M.I., Marks M.S.;
RT "Characterization of mouse tGolgin-1 (golgin-245/trans-golgi p230/256
RL KD golgin) and its upregulation during oligodendrocyte development.";
RN DNA Cell Biol. 21:505-517(2002).
RN [2]
RP SEQUENCE OF 1-980 AND 991-2231 FROM N.A., AND VARIANTS ARG-61;
RP SER-280; GLU-293; SER-638; THR-819; THR-829 AND ALA-859.
RC STRAIN=C57BL/6, and Czech II; TISSUE=Brain, and Breast cancer;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Distchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May play a role in vesicular transport from the trans-
CC Golgi (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic, peripheral membrane protein
CC associated with the trans-Golgi network.
CC -!- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in
CC oligodendrocyte precursors, particularly at a stage just prior to
CC myelination.
CC -!- DOMAIN: Extended rod-like protein with coiled-coil domains.
CC -!- SIMILARITY: Belongs to the golgin family.
CC -!- SIMILARITY: Contains 1 GRIP domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DQ EMBL; AF051357; AAC05573.2; -.
DQ EMBL; BC007485; AAH07485.1; -.
DQ EMBL; BC037641; AAH37641.1; ALT_INIT.
DQ EMBL; BC053000; AAH53000.1; -.
DQ PIR; T14265; T14265.
DQ MGI; MGI:1859646; Golga4.
DQ GO; GO:0005794; C:Golgi apparatus; IDA.
DQ InterPro; IPR002377; GRIP_Domain.
DQ InterPro; IPR003345; M_repeat.
DQ Pfam; PF01465; GRIP; 1.
DQ Pfam; PF02370; M; 12.
DQ PROSITE; PS05913; GRIP; 1.
KW Golgi stack; Coiled coil; Polymorphism.
DQ DOMAIN 156 2161 COILED COIL (POTENTIAL).
```

FT	DOMAIN	335	1011	GLU-RICH.
FT	DOMAIN	1855	2107	GLU-RICH.
FT	DOMAIN	2178	2225	GRIP.
FT	VARIANT	61	61	T -> R (in strain Czech II).
FT	VARIANT	280	280	G -> S (in strain Czech II).
FT	VARIANT	293	293	G -> E (in strain Czech II).
FT	VARIANT	638	638	G -> S (in strain Czech II).
FT	VARIANT	819	819	K -> T (in strain Czech II).
FT	VARIANT	829	829	A -> T (in strain Czech II).
FT	VARIANT	859	859	T -> A (in strain Czech II).
FT	VARIANT	2065	2065	R -> S (in strain C57BL/6).
FT	SEQUENCE	2238 AA; 257562 MW; 494EA2C111F0165B CRC64;		
Query Match				
Best Local Similarity 19.5%; Pred. No. 4.9e-11;				
Matches 419; Conservative 321; Mismatches 612; Indels 799; Gaps 90				
QY	4	FKYGARNPLDAGAAPIASRASRLNL	-----FFQCKPPMTQQQMSPLSREGILD	53
DB	98	FRSPIKESLFRSSKEPIVRTSGRESLNQLDLCSAAFDPSPDMESEADAPNSDGLSR	157	
QY	54	ALFVLFECSQPALMKIKHVNFRVKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAVQV	113	
DB	158	-----EQLLQRLRMERSLSYRGKYSELVTFQTLQ	-----	189
QY	114	VREKATGIYAMKVMKKALLAQOVFFEEERNILSRSTSPWIPOLQYAFQDKXHLYL	173	
DB	190	-REK-----KKLQILUSQO	-----DKSL--RRISEURELQWQQAQKHL	227
QY	174	BEYQPGDLLSLNRYEDQLDENLQFYLAELILAVHSGVHMDIKPENILVDRGT	233	
DB	228	-----QDEPDAFL	-----EEKQYVISLQIT	247
QY	234	HLKLVDFGSAAMNSKMNNAKLPIGTEDYMAP	-----EVLTVNNGDGKTYGLDCDWSVGVI	292
DB	248	-----QVSLKQLQLONGPMNVDPAPKLPPGELQAEVHGTEKNEG	-----VG--	289
QY	293	AYEMTYGRSPAEGTSARTFNIMNFRFLKFPDPKVSDFDLQ	-----	339
DB	290	-----EPVGGTISAKTLEML	-----QQRVKRQE-----NLLQRCKEITGSHKQC	329
QY	340	SLLCQKRELKPEGLCCHPFPSKIDWNINRSPPPVPTLKSDDDTSNFDPEPKNSWSS	399	
DB	330	ALLLSEKEALQ	-----EQJD-----ERLQELERNKELHMAEKTCLIT	366
QY	400	SPCQLSPGSGSEELPFVGFYSYKALGILGRSESVYSGL	-----DSPAKTSSMEKLLTKS	455
DB	367	---QLRDKNLIEQL	-----EQDKGMVITTKQMLETLEKDETAQLRSHIKOMTQG	418
QY	456	KELODSQK	-----CHKQEQMTLHRRVSEBAVLQSQKEVELKASQTR	500
DB	419	EBLREQEKSERAAFELEKALSTAQKTEDAQRMKMEMDE	-----QMKAVE-RASEBER	472
QY	501	SLLQEDLATYITECSSLKSRLSEQARMEVSQDDKALQLLH	-----	540
DB	473	LRQLQELSRVQEAASNAK	-----KNSBEQVAALQKHAEELAKQBELSRRLEABER	525
QY	541	DIREQSR	-----KLQEIKEQYQYQVEEMRL	585
DB	526	ELQEQMRIALESRSRYLKLQTEKEQESLASELELQKKAILTSENKLQELQGEAEAY	585	
QY	586	ESELRESLAAEE-FYRKATCCQH	-----KLLKAKDGCKPEVCEY	624
DB	586	RTRILETSTLEKSLQESQTQSEHLAVHLEAKNKNKINKELTALAEQHRTEVEGLQQQDS	645	
QY	625	---AKLEKINAEQQLKTELQEKLI	-----EKAA--KERA-----EREKELQNREDS	666
DB	646	LWTEFQSLSQHQAAVEELREKYQOEKDALLEKESLQCAHTQDMNKTLEKLDKKQWE	705	
QY	667	SEGRTKKLVEA	-----EERRH	682
DB	706	LESVSSSELSELRARDQLAEELSILRGDADKMKQALEAELESPQRHHRHREVGSI	765	

Db 1809 ----CLPEELEBKCKSLIVSPMG-----EETGNTGVK-----Q 1840

Qy 1474 GNMKVPNNKRGQGHXRVYIVLEGSKVLYDNEAREAGQVSEFELCLPDGVSIHCA 1533

Db 1841 NNASVVDVSQKTLQEKELTCQALEQRVKEESDLVRERGAHLEVEKL----- 1888

Qy 1534 VGASELANTAKAEKARADAKLAGNSLLKLEGDDRLDMNCTLPFSQVVLVGTGEGLYALN 1593

Db 1889 -----TLKYEKSQ-----SSQEMDGENK-----CVELEDR-----FEENSQSHE 1924

Qy 1594 VLKNSLTHVPGGAVFQIIVIIKDLEKLMIAGEERALCLVDVKKVKQSIAQ 1644

Db 1925 IQSNVGT-VDGLRS-----DLESKL--TGAER-----DQKLSKEVAR 1959

RESULT 10

CEP2\_HUMAN

ID CEP2\_HUMAN STANDARD; PRT; 2442 AA.

AC Q9BV73; O14812; O60588; Q9H450;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Centrosomal protein 2 (Centrosomal Nek2-associated protein 1) (C-NAP1)

DE Centrosome protein 250 (Centrosome associated protein CEP250).

GN CEP2 OR CNAP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1] CEP2\_HUMAN

RP SEQUENCE FROM N.A. (ISOFORM 1), AND AUTOIMMUNE DISEASE.

RC TISSUE=Cervical carcinoma;

RX MEDLINE=98165428; PubMed=9506584;

RA Mack G.J., Rees J., Sandblom O., Balczon R., Fritzler M.J.,

RT "Autoantibodies to a group of centrosomal proteins in human autoimmune

RT sera reactive with the centrosome."

RL Arthritis Rheum. 41:551-558(1998).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), INTERACTION WITH NEK2,

RP AND SUBCELLULAR LOCATION DURING THE CELL CYCLE.

RC TISSUE=Placenta;

RX MEDLINE=98311641; PubMed=9647649;

RA Fry A.M., Mayor T., Meraldi P., Stierhof V.-D., Tanaka K., Nigg E.A.;

RT "C-Nap1, a novel centrosomal coiled-coil protein and candidate

RT substrate of the cell cycle-regulated protein kinase Nek2."

RL J. Cell Biol. 141:1563-1574(1998).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagdley C.L.,

RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,

RA Lhvasialho M.H., Ievers W.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConnachie L.J., McIay K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,

RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,

RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,

RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RA Rogers J.;

RL "The DNA sequence and comparative analysis of human chromosome 20.";

RN Nature 414:865-871(2001).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 3).

RC TISSUE=Placenta;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.I., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,

RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [5]

RP PHOSPHORYLATION, AND INTERACTION WITH NEK2 AND PPP1CA.

RX MEDLINE=21060765; PubMed=10880350;

RA Helps N.R., Luo X., Barker H.M., Cohen P.T.W.;

RT "NIMA-related kinase 2 (Nek2), a cell-cycle-regulated protein kinase

RT localized to centrosomes, is complexed to protein phosphatase 1."

RL Biochem. J. 349:509-518(2000).

RN [6]

RP PHOSPHORYLATION DURING CELL CYCLE.

RX MEDLINE=22135747; PubMed=12140259;

RA Mayor T., Hacker U., Stierhof V.-D., Nigg E.A.;

RT "The mechanism regulating the dissociation of the centrosomal protein

RT C-Nap1 from mitotic spindle poles."

RL J. Cell Sci. 115:3275-3284(2002).

CC -!- FUNCTION: Probably plays an important role in centrosome cohesion

CC during interphase.

CC -!- SUBUNIT: Monomer and homodimer (Probable). Forms a complex in

CC vitro with both NEK2 kinase and the PPP1CC catalytic subunit of

CC protein phosphatase 1 (PP1).

CC -!- SUBCELLULAR LOCATION: Component of the core centrosome. In

CC interphase cells, it specifically associates with the proximal

CC ends of both mother and daughter centrioles. Associates with the

CC centrosome in interphase cells. In mitotic cells, it dissociates

CC from the mitotic spindle poles. At the end of cell division, it

CC reaccumulates at centrosomes.

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=3;

CC Name=1;

CC IsoId=Q9BV73-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q9BV73-2; Sequence=VSP\_007372;

CC Name=3;

CC IsoId=Q9BV73-3; Sequence=VSP\_007370, VSP\_007371;

CC Note=No experimental confirmation available;

CC -!- TISSUE SPECIFICITY: Ubiquitously and weakly expressed.

CC Note=PTM: Differentially phosphorylated during cell cycle.

CC Phosphorylation may regulate association/dissociation from

CC centrosome. During M phase of mitosis, C-terminal part is

CC phosphorylated by NEK2, suggesting that it may trigger the

CC dissociation from the mitotic centrosome. It is dephosphorylated

CC in vitro by the PP1 phosphatase.

CC -!- DISEASE: Antibodies against CEP2 are present in sera from patients

CC with autoimmune diseases that developed autoantibodies against

CC centrosomal proteins.







split into 2 globular subfragments (S1) and 1 rod-shaped subfragment (S2).

-!- DOMAIN: The rodlike tail sequence is highly repetitive, showing cycles of a 28-residue repeat pattern composed of 4 heptapeptides, characteristic for alpha-helical coiled coils.

-!- PTM: Phosphorylation inhibits thick filament formation and reduces the actin-activated ATPase activity.

-!- MISCELLANEOUS: Dictyostelium myosin II has no K(2)EDTA ATPase activity, perhaps correlated with the absence of a Cys at the SH-1 position (588).

-!- SIMILARITY: Contains 1 myosin-like globular head domain.

-!- SIMILARITY: Contains 1 IQ domain.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; M14628; AAA33227.1; -.

PIR; A26655; A26655.

PDB; 1MVA; 03-DEC-97.

PDB; 1MMD; 17-AUG-96.

PDB; 1MMG; 03-DEC-97.

PDB; 1MNN; 03-DEC-97.

PDB; 1MND; 17-AUG-96.

PDB; 1MNE; 17-AUG-96.

PDB; 1VOM; 23-DEC-96.

PDB; 1LVK; 28-JAN-98.

PDB; 1DOX; 20-DEC-00.

PDB; 1DOY; 20-DEC-00.

PDB; 1DOZ; 20-DEC-00.

PDB; 1D1A; 20-DEC-00.

PDB; 1D1B; 20-DEC-00.

PDB; 1D1C; 20-DEC-00.

PDB; 1FMV; 20-DEC-00.

PDB; 1G8X; 17-JAN-01.

PDB; 1JWY; 07-NOV-01.

PDB; 1JX2; 07-NOV-01.

DictyBase; DDB002015; mhca.

InterPro; IPR000048; IQ region.

PRINTS; PR00193; Myosin\_N.

InterPro; IPR004009; Myosin\_N.

InterPro; IPR008989; Myosin\_S1\_N.

Pfam; PF00612; IQ; 2.

Pfam; PF00683; myosin\_head; 1.

Pfam; PF02736; Myosin\_N; 1.

PRINTS; PR00193; MYOSINHEAVY.

ProDom; PD000355; myosin\_head; 1.

SMART; SM00015; IQ; 1.

SMART; SM00242; MYSC; 1.

PROSITE; PS50096; IQ; 1.

Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure; Calmodulin-binding; Methylation; Phosphorylation.

FT DOMAIN 1 761 MYOSIN HEAD-LIKE.

FT DOMAIN 762 791 IQ.

FT DOMAIN 817 2116 COILED COIL (POTENTIAL).

FT NP\_BIND 179 186 ATP.

FT DOMAIN 638 660 ACTIN-BINDING.

FT DOMAIN 738 752 ACTIN-BINDING.

FT MOD\_RES 130 130 METHYLATION (DI-) (POTENTIAL).

FT MOD\_RES 1823 1823 PHOSPHORYLATION (BY MECK).

FT MOD\_RES 1833 1833 PHOSPHORYLATION (BY MECK).

FT MOD\_RES 2029 2029 PHOSPHORYLATION (BY MECK).

FT TURN 3 5

FT TURN 7 8

FT HELIX 10 15

FT HELIX 25 28

FT STRAND 34 37

FT TURN 41 42

FT STRAND 48 55

FT STRAND 59 63

FT STRAND 69 73

FT HELIX 74 76

FT STRAND 78 79

FT HELIX 83 85

FT TURN 86 87

FT STRAND 90 90

FT HELIX 91 93

FT HELIX 99 110

FT TURN 111 113

FT STRAND 116 119

FT TURN 120 121

FT STRAND 122 126

FT HELIX 137 142

FT TURN 143 145

FT HELIX 148 150

FT HELIX 155 169

FT TURN 170 170

FT STRAND 173 179

FT TURN 181 182

FT HELIX 185 200

FT HELIX 210 226

FT STRAND 227 228

FT TURN 234 235

FT STRAND 236 237

FT STRAND 240 247

FT TURN 249 250

FT STRAND 253 261

FT HELIX 265 268

FT TURN 269 269

FT TURN 273 274

FT STRAND 278 278

FT HELIX 279 287

FT TURN 290 296

FT TURN 297 297

FT HELIX 301 303

FT TURN 305 307

FT TURN 316 317

FT HELIX 320 334

FT TURN 335 335

FT HELIX 338 355

FT TURN 356 356

FT STRAND 360 360

FT STRAND 368 368

FT HELIX 373 382

Query Match 5.2%; Score 540.5; DB 1; Length 2116;

Best Local Similarity 21.5%; Pred. No. 5.5e-10;

Matches 326; Conservative 281; Mismatches 505; Indels 405; Gaps 68;

QY 58 LPEECQPALMKIKHVSNFVRKYSYDTIAELQELQPSAKDFEVRSLVGCQGEFA-----EVQ 112

DB 534 VFNATDNTLITKLH-SHFSKQN-----AKYEPRFKTEF-----GVTHYAGQVMYEIQ 582

QY 113 VVEKATGDIYAMKVMKKALLAQEQVSFFEEERNILSRSTSPWIFQLQYAFQDKHLYL 172

DB 583 DWLEK-----NKDPLQODLELCFKSDSNVVK-----LFNDPN---I 617

QY 173 MEELYQGGDLLSLNRYEQDLNLIQFVLAELILAVSHVLMGYVHRDIKPEN----- 226

DB 618 ASRAKGANFITVAAQYKEQ-----LASLMATLETTN--PHFVRCIIPNNKQLPAK 666

QY 227 -----ILVDR---TGHIKLVDFGSAAKNNKMNNAKLPIGTDPY-----MAPEVLTVNN 273

DB 667 LEDKVVLDQLRCNGVLEGIRI--TRKGFNRIYA-----DFVKRYVLLAPNVPDAE 717

QY 274 GDGKGT-----YGLDCDWSVGVIAYEIMYGRSPFAEGTSARTFNNINMFORFLKFPDD 327

DB 718 DSQKATDAVLKHLNIDPEQYRFGIT--KIFFRAGQLARIEEARE-----QRI----- 762

QY 328 PKVSSDFLDLIQSLCGQXERLKFEGLCCHPFFSKIDWNIR-----NSPPPVV----- 376



Db 119 DFEILNLVGKSGFGKVIQVRKKDGTGEVYAMKVLKKHIVEHNEVHTLSERNILQKINHP 178  
 Qy 156 WIPOLQYAFQDKNHLVMEYQGGDLNLRVEDQDLENLIQFYLAELILAVHSHLM 215  
 Db 179 FLVNLNYSFQTEKLYFILDYVNGGELFVHLQK-DKKFTEDRVRYGAEIVLALHLS 237  
 Qy 216 GYVHRDIPENILVDRTHIKLVDFGSAKNSKNVAKPIGTPDYMAPEVLTVNGD 275  
 Db 238 GVIYEDLPENILLTNEGHCMTDFGLCKEGLLTPTDKTGFQCPPEYLAPEVL---QGN 294  
 Qy 276 KGSTYGLDGDMSVGVIAVEMYIGRSPFAEGTSARTFNINMFQFLFPDPPKVSDFL 335  
 Db 295 G---YGVQVDSWFSGLLYEMLTGLPPFYNQDVQEMVRKIM-MEKLSPFH--FISDPAR 347  
 Qy 336 DLIQSLCCQKE-RUKFGLC-CHFFPSKIDWNI--RNSPPFPVPTLKSDDTSNFD-- 389  
 Db 348 SLLEQLERDPKRLADPNLIKRPFFRSIDWEQLFQKNIPPPFPNVKGSADTSQIDV 407  
 Qy 390 -EPEKNSVSSPCCOLSPSGSGEPLPVGFSY 421  
 Db 408 FDEAPSLTMAGECALNIQ-----QQKDFEGFTY 436

RESULT 14  
 KAIB SCHPO  
 ID KAIB SCHPO STANDARD; PRT; 607 AA.  
 AC Q09598.  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Probable serine/threonine-protein kinase C24B1.1lc (BC 2.7.1.-).  
 GN SPAC24B1.1lc.  
 OS Schizosaccharomyces pombe (Fission Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor A., Cronin A., Davis P., Feltwell J., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leacher S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potaashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe";  
 RL Nature 415:871-880(2002).  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC HIGHLY SIMILAR TO YEAST DBF20 AND DBF2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: Z67757; CAA91776.1; --  
 DR PIR: S62556; S62556.  
 DR GeneD\_Sfombe; SPAC24B1.1lc; --  
 DR InterPro; IPR000961; Pkinase\_C.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00433; pkinase\_C; 1.  
 DR ProDom; EDD00001; Prot\_Kinase; 2.  
 DR SMART; SMO0133; S\_TK\_X; 1.  
 DR SMART; SMO0220; S\_TK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DM; 1.  
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;  
 KW ATP-binding; Phosphorylation.  
 FT DOMAIN 208 508 PROTEIN KINASE.  
 FT NP\_BIND 214 222 ATP (BY SIMILARITY).  
 FT BINDING 237 237 ATP (BY SIMILARITY).  
 FT ACT\_SITE 331 331 BY SIMILARITY.  
 FT MOD\_RES 219 219 PHOSPHORYLATION (BY SIMILARITY).  
 SQ SEQUENCE 607 AA; 70491 MW; D9285263C7341FID CRC64;  
 Query Match 5.1%; Score 540; DB 1; Length 607;  
 Best Local Similarity 28.6%; Pred.No.1.7e-10;  
 Matches 142; Conservative 85; Mismatches 150; Indels 120; Gaps 15;  
 Qy 27 LNLFFQKQPFMTQQQMSPLSREGILDALFVLFEECSQPALMKIKHVSNFVRKYSD--T 83  
 Db 149 LDHYFEQLHYLYTRKQARLFEEQL-----KEPDSRR-----DELVKRYNGRERV 194  
 Qy 84 IAEQLQPSAKDFEVRSLVCGGHFAEVQVVRKATGDIYAMKVMKKALLAQEQSF 143  
 Db 195 YLKRRTTRISHGDFQITQVGGYGVSWLARKRDTKEI VALKIMKSVLHKMDEIRHVL 254  
 Qy 144 BERNILSRSTSPWIPQLOAYAFQDKNHLVMEYQGGDLNLRVEDQDLENLIQFYLA 203  
 Db 255 TERDILTANSEWLVRLLYAFQDTSNIYLA MEYVPGDFTLLSN-SGVLRDHAKFYAT 313  
 Qy 204 ELILAVHSHLMGVHRDIKPNILVDRTHIKLVDFGSAKNSKNV----- 252  
 Db 314 EMFLADALHQLGYIHRDLKPNFLVGCASHIKLTFGLSSGIISKXIESMKIRLQEVN 373  
 Qy 253 NAKLP-----IGTPDYMAPEVLTVMGDGGKGTGVLDCDWM 287  
 Db 374 NVVVPERSMERRQVFRLLSQDPVVAHSVGVSPDYMAPEVLRGEN-----YNHSVDYW 427  
 Qy 288 SVGVIAVEMYIGRSPFAEGTSARTFNINMFQFLKFP--DDPK-----VSSDFLDLIQS 340  
 Db 428 SLGCMYTELUSGPPFPGSGSNVNTWSNLKNRKCFCORPHYDDPRDLFEFNWRDDAWEVCH 487  
 Qy 341 LLCQKRL-KFEGLCCHPFFSKIDWNIENS-PPFPVPTLKS-----DDTSN----- 387  
 Db 488 CIITDPKDFCSLKQVMQHPYFSKIDWNVRTAYRPPFVPLNGLNGLNGLNGLNGLNGLN 547  
 Qy 388 -----FDEPKNSVSSPCCOLSPSGSGEPLPVGFSYKALGILG 429  
 Db 548 KYKEVHEKQAAIANMVNTFNKPKRNA-----PIGTFRRHQKNSHP 587  
 Qy 430 RRSRVSGLDSPAKTS 446  
 Db 588 TSSS--SALSSPLSAPS 602  
 RESULT 15  
 MYHA RAT  
 ID MYHA\_RAT STANDARD; PRT; 1976 AA.

AC Q9JLT0; 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Myosin heavy chain, nonmuscle type B (Cellular myosin heavy chain,  
DE type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).  
GN MYH10  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
RX MEDLINE=20483650; PubMed=11027611;  
RA Yam J.W.P., Chan K.W., Li N., Hsiao W.L.W.;  
RT "Molecular cloning and functional analysis of the promoter region of  
RT rat nonmuscle myosin heavy chain-B gene."  
RL Biochem. Biophys. Res. Commun. 276:1203-1209(2000).  
CC -!- FUNCTION: Cellular myosin appears to play a role in cytokinesis,  
CC cell shape, and specialized functions such as secretion and  
CC capping.  
CC -!- SUBUNIT: Myosin is a hexameric protein that consists of 2 heavy  
CC chain subunits (MHC), 2 alkali light chain subunits (MLC) and 2  
CC regulatory light chain subunits (MLC-2).  
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing  
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,  
CC characteristic for alpha-helical coiled coils.  
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.  
CC -!- SIMILARITY: Contains 1 IQ domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AFI39055; AAF61445.1; --  
DR HSP; P10587; 13R2.  
DR InterPro; IPR000048; IQ\_region.  
DR InterPro; IPR001609; myosin\_head.  
DR InterPro; IPR004009; Myosin\_N.  
DR InterPro; IPR002928; Myosin\_tail.  
DR InterPro; IPR002017; Spectrin.  
DR Pfam; PF00612; IQ; 1.  
DR Pfam; PF00063; myosin\_head; 1.  
DR Pfam; PF02736; Myosin\_N; 1.  
DR Pfam; PF01576; Myosin\_tail; 1.  
DR PRINTS; PD00193; MYOSINHEAVY.  
DR ProDom; PD000355; myosin\_head; 1.  
DR SMART; SM00015; IQ; 1.  
DR SMART; SM00242; MYSC; 1.  
DR PROSITE; PS50056; IQ; 1.  
DR PROSITE; PS50056; IQ; 1.  
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding; Coiled coil;  
KW Multigene family.  
KW  
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.  
FT DOMAIN 786 815 IQ.  
FT DOMAIN 845 1976 COILED COIL (POTENTIAL).  
FT NP\_BIND 178 185 ATP (POTENTIAL).  
SQ SEQUENCE 1976 AA; 228963 MW; E32708BF9BF2B470 CRC64;  
  
Query Match 5.1%; Score 538; DB 1; Length 1976;  
Best Local Similarity 23.2%; Pred. No. 6.2e-10;  
Matches 257; Conservative 218; Mismatches 355; Indels 279; Gaps 47;  
  
QY 443 KTSSMEKLLIKSLQSDQKCHME---QSMTRLHRRVSEAVLS---QKEVELKA- 495  
DB 846 QYTRQEEELQAKDELLKVKYKTKVGELEEMERKHQQLLEKNILAEQLAETELFAE 905  
  
QY 496 QYTRQELL---EODLATYITESSLSRLEQARMEVSQDDKALQI-LHDREQRKLOE 551  
DB 496 QYTRQELL---EODLATYITESSLSRLEQARMEVSQDDKALQI-LHDREQRKLOE 551

DB 906 AEEMEARLAACKQLEBEILHDLSSRVGE--BEERNQILQNEKKWQAHIQDLBEQLDE-EE 963  
QY 552 IKEEYQAQVEEVRMLMNQIUED--LVSARRRDLSELSRSLAAEFKATCEQHK 609  
DB 964 GARQKQLEKVTABAKIKWEEVLLLEDQNSKFKEKKLMEDRIA-----ECSSQ 1014  
QY 610 LLKAKDGKPEVGEYAKLEKINAEQOLKIOELOKLEKAAK-----ERARELE- 658  
DB 1015 LAEEERAK-----NLAKIRNKQEVMSIDLEERLJKEKTRQELEKAKKLDGTTDL 1067  
QY 659 -----KLNREDSSG-----IRKKLVAREE-- 679  
DB 1068 QDQIAELQAQVDELKVLTKKEBELQALARGDDETLHKNNALKVARELQAQIAELQEDF 1127  
QY 680 -----RHSLENKVKEL-ETMERRENRLKDDIOT-----KSQIQOMAD--KILELEE 724  
DB 1128 ESEKASRNKAEKQKRDLSBELEALKTELEDTLDTAAQBELTKRQEVAEALKKALEDET 1187  
QY 725 KHREAAQVS-----AQHLEVLHXQEHYBEKIKVLDNQIKKDLADKE---TLNMMQRH 775  
DB 1188 KNHEAQ-QDMRQRHATALEBELSQLEQAKRFKANLEKKNQCOLETDNKLACACEVKVLQVVK 1247  
QY 776 EEEAEKKGKILSEKAMINAMDSKIRSLEQRIVELSE-ANKLAANSSLFTORMKQAEEM 834  
DB 1248 ABSEHKKKL-----DAQVQLHAKVSGDRLRVELAEKANKL-----QNELDNVSTL 1295  
QY 835 ISELROQKPFYLETQAGKLEAQNRLKEBELEKISHQDHSKNNRLLELE---TRLREVSLEH 891  
DB 1296 LEEAEKGMKFAKDAAGLESQLODTQELLQOEETRQKLNLSRIRLEEEKNSLQEQEEE 1355  
QY 892 BEQKLELKRQJTELQLSQERESQ-----LTAQAAR-----AALESQRQ--AK 934  
DB 1356 BEARKNLEKVLQSLQSLADTKKKVDDDLGTIEGLEAEKKKLDKVEALSQRLEEKVLAY 1415  
QY 935 TELBETTAEBEETQALTA--HRDEI-----QRKFD-ALRNSCTVITDLEEQNLQT 983  
DB 1416 DKLEKTKNRLOEELDDLTVDLHQRIQVSNLEKKQKFDOLLAEEKISARYAEERDRAE 1475  
QY 984 EDNAELNNQNFYISKQLEDEASGANDEI---VQJRSEV-----DHLRREITEREMQLT 1032  
DB 1476 AEAREKETKALSARALEEALEAEKESFERQNKQLRADMEDLMSSKDDVGNVHELEKSR 1535  
QY 1033 SQKQTMALKTTCTMLEEQVMDLEALNDELLEKERQWEAWRSVL-----GDEKSQFECH 1086  
DB 1536 ALEQQVBEEMR---TQLEELDELEIQAEDAKLRLEVMNQAKAOFERDLQTRDQNEKFR 1592  
QY 1087 -----VRELQRLMDTEKQSPARADQRTESRQVVELAVKEHKAELIA----- 1128  
DB 1593 LLLKQVRELEAELEDERKQALAA---VASKKKWEIDLKLEAQIEAANKARDEVIKQLR 1648  
QY 1129 -----LQALKEQKLKAESLSKLNDELKXHAMLENNARSLQKL-----ET 1170  
DB 1649 KLQAKMDYQRELEEARASDEIFAQSEKSEKLSLEAILQLEBELASSERARHAEQ 1708  
QY 1171 ER-EL-----KQRLLEEQ-----AKLQQQMDLQKNHIFRITQGLQBALDEADL 1212  
DB 1709 ERDELADEIANSAGKSALLDEKRRLEARIAQLEELSEEBEQSNWELNDRFRKTTIQLVDT 1768  
QY 1213 LKTE-----RSDLEYQLE-----NTQVLYSHKVKVMEGTISQOTKLIDFLQA 1254  
DB 1769 LNTLEAASRAAKSDNARQQLERONKELKAKUQLEAGVSKFKATISALEAKIQLEE 1828  
QY 1255 KMDQPA-----KXKGLF-----SRRKEDPALPTQVPLQYNELKALEKEK 1295  
DB 1829 QLEGEAKERAANKLVPRTEKKLEIFMVEDDERHAD-----QYKE---QMEKAN 1876  
QY 1296 ARCAELBEALQKTRIELRSAREEAAHRK 1323  
DB 1877 ARMKQLKQLEEA--EBEATRANASRRK 1902

Search completed: July 3, 2004, 10:17:57  
Job time : 27 secs



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:12:55 ; Search time 27 Seconds  
(without alignments)  
3925.486 Million cell updates/sec

Title: US-10-017-216-2

Perfect score: 10450  
Sequence: 1 MLKFKYGARNPLDAGAAEPI.....QLNGEIROQVEKSVLRTDYC 2053

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%  
Maximum Watch 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2430	23.2	497	4	US-09-804-471A-2
2	2430	23.2	497	4	US-10-238-709-2
3	2165.5	20.6	494	4	US-09-804-471A-4
4	2165.5	20.6	494	4	US-10-238-709-4
5	1277	12.2	257	4	US-09-916-204-2
6	1253	11.9	1388	2	US-08-685-576-1
7	1250.5	11.9	1354	3	US-08-685-871-2
8	1249.5	11.9	1388	4	US-09-976-594-296
9	1247.5	11.9	1388	2	US-08-685-576-4
10	1173	11.2	251	4	US-09-916-204-4
11	1170	11.2	251	4	US-09-916-204-5
12	1170	11.2	251	4	US-09-916-204-6
13	1043.5	9.9	900	2	US-08-630-822A-62
14	1043.5	9.9	900	2	US-09-005-063-62
15	1043.5	9.9	900	4	US-09-171-156A-21
16	1043.5	9.9	900	4	US-09-004-730A-21
17	1043.5	9.9	900	4	US-08-981-799A-21
18	847.5	8.1	420	3	US-08-685-871-58
19	843.5	8.0	582	2	US-08-422-699A-9
20	843.5	8.0	582	2	US-08-422-706B-9
21	830.5	7.9	420	3	US-08-685-871-59
22	799.5	7.6	638	2	US-08-422-706A-11
23	799.5	7.6	638	2	US-08-422-706B-11
24	776	7.4	555	1	US-08-484-044-6
25	706.5	6.7	479	4	US-09-442-100-13
26	706.5	6.7	479	4	US-08-939-106-13
27	706.5	6.7	479	4	US-09-442-102-13

28	691	6.6	526	4	US-09-442-100-12	Sequence 12, Appl
29	691	6.6	526	4	US-08-939-106-12	Sequence 12, Appl
30	691	6.6	526	4	US-09-442-102-12	Sequence 12, Appl
31	677.5	6.5	719	3	US-09-588-256-2	Sequence 2, Appl
32	673	6.4	404	2	US-08-860-150-3	Sequence 3, Appl
33	673	6.4	404	3	US-09-338-132-3	Sequence 3, Appl
34	668	6.4	464	2	US-08-878-989-4	Sequence 4, Appl
35	668	6.4	464	3	US-09-272-796-4	Sequence 4, Appl
36	665.5	6.3	465	2	US-08-878-989-18	Sequence 18, Appl
37	665.5	6.3	465	2	US-08-860-150-7	Sequence 7, Appl
38	665.5	6.3	465	3	US-09-338-132-7	Sequence 7, Appl
39	665.5	6.3	465	3	US-09-272-796-18	Sequence 18, Appl
40	661	6.3	500	4	US-09-442-100-14	Sequence 14, Appl
41	661	6.3	500	4	US-08-939-106-14	Sequence 14, Appl
42	661	6.3	500	4	US-09-442-102-14	Sequence 14, Appl
43	630	6.0	620	4	US-09-442-100-11	Sequence 11, Appl
44	630	6.0	620	4	US-08-939-106-11	Sequence 11, Appl
45	630	6.0	620	4	US-09-442-102-11	Sequence 11, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-804-471A-2  
; Sequence 2, Application US/09804471A  
; Patent No. 6479269  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CLO01164  
; CURRENT APPLICATION NUMBER: US/09/804.471A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 497  
; TYPE: PRT  
; ORGANISM: Human  
US-09-804-471A-2

Query Match	23.2%	Score 2430;	DB 4;	Length 497;
Best Local Similarity	99.1%	Pred. No. 8.8e-118;		
Matches	464;	Conservative	3;	Mismatches 1; Indels 0; Gaps 0;
QY	1	MLKFKYGARNPLDAGAAEPIASRASRLNLFQKPPMTQQMSPLSREGILDALFVLFE	60	
DB	1	MLKFKYGARNPLDAGAAEPIANRASRLNLFQKPPMTQQMSPLSREGILDALFVLFE	60	
QY	61	ECSPALMKIKHVSFVRKYSDDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG	120	
DB	61	ECSPALMKIKHVSFVRKYSDDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG	120	
QY	121	DIYAMKVMKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKXNHLVMEYQFGG	180	
DB	121	DIYAMKVMKKALLAQEQVSFFEEERNILSRSTSPWIPQLQYAFQDKXNHLVMEYQFGG	180	
QY	181	LLLSLLNRYEDQDENLIQFYLAELIILAVHSVHLMGVVHRDIPKENTILVDRGTGKILVDF	240	
DB	181	LLLSLLNRYEDQDENLIQFYLAELIILAVHSVHLMGVVHRDIPKENTILVDRGTGKILVDF	240	
QY	241	GSAAMNSKNVNAKLPIGTPDYNAPEVLTVMNGDGKGTGYGLDCDMSVGVIAEYMIYGR	300	
DB	241	GSAAMNSKNVNAKLPIGTPDYNAPEVLTVMNGDGKGTGYGLDCDMSVGVIAEYMIYGR	300	
QY	301	SPFAEGTSARTFNIMNFQRLKFPDPKVSSDFLDLIQSLLCQCKERLKFEGJCCHPFF	360	
DB	301	SPFAEGTSARTFNIMNFQRLKFPDPKVSSDFLDLIQSLLCQCKERLKFEGJCCHPFF	360	
QY	361	SKIDWNIRSPFPVPTLKSDDDTSNFDEPKNSWVSSPCLSPSGFSGEELPFYGF	420	

```
Db 361 SKIDWNIRNSPPFVPTLSDDDTSNFDPEKNSWVSSSPCQLSPSGFSGEELPFVGF 420
Qy 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQSDQKCHM 468
Db 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQSDQKCHV 468

RESULT 2
US-10-238-709-2
; Sequence 2, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Human
US-10-238-709-2

Query Match 23.2%; Score 2430; DB 4; Length 497;
Best Local Similarity 99.1%; Pred. No. 8.8e-118;
Matches 464; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLKFYKGYARNPLDAGAAEPIASRASRLNLPFGKPPMTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFYKGYARNPLDAGAAEPIASRASRLNLPFGKPPMTQQQMSPLSREGILDALFVLFE 60

Qy 61 ECSOPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 120
Db 61 ECSOPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 120

Qy 121 DIAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLOQYAFQDKNHLVMEYOPGG 180
Db 121 DIAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLOQYAFQDKNHLVMEYOPGG 180

Qy 181 DLSLLNRYEDQDENLIQFYLAELIIVAVSHVLMGVVHRDIKPNILVDRTHIKLVDF 240
Db 181 DLSLLNRYEDQDENLIQFYLAELIIVAVSHVLMGVVHRDIKPNILVDRTHIKLVDF 240

Qy 241 GSAAKMNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKGTGLDCDMSVGVIAIYEMIYGR 300
Db 241 GSAAKMNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKGTGLDCDMSVGVIAIYEMIYGR 300

Qy 301 SPFAEGTSARTFNINMFQRLKFPDPPKYSDFLDLIQSLLCQKRLKFEGLCCHPFF 360
Db 301 SPFAEGTSARTFNINMFQRLKFPDPPKYSDFLDLIQSLLCQKRLKFEGLCCHPFF 360

Qy 361 SKIDWNIRNSPPFVPTLSDDDTSNFDPEKNSWVSSSPCQLSPSGFSGEELPFVGF 420
Db 361 SKIDWNIRNSPPFVPTLSDDDTSNFDPEKNSWVSSSPCQLSPSGFSGEELPFVGF 420

Qy 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQSDQKCHM 468
Db 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQSDQKCHV 468

RESULT 3
US-09-804-471A-4
; Sequence 4, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
```

```
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Mus Musculus
US-09-804-471A-4

Query Match 20.6%; Score 2165.5; DB 4; Length 494;
Best Local Similarity 87.8%; Pred. No. 3.6e-104;
Matches 411; Conservative 26; Mismatches 30; Indels 1; Gaps 1;

Qy 1 MLKFYKGYARNPLDAGAAEPIASRASRLNLPFGKPPMTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFYKGYARNPEASASEPIASRASRLNLPFGKPPMTQQQMSALSREGMLDALFALFE 60

Qy 61 ECSOPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVRSLVCGGHFAEVQVVRKATG 120
Db 61 ECSOPALMKIKHVSFVQKYSDTIAELRELOPSARDFEVRSLVCGGHFAEVQVVRKATG 120

Qy 121 DIAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLOQYAFQDKNHLVMEYOPGG 180
Db 121 DIAMKVMKKKALLAQEQVSFFEEERNILSRSTSPWIPQLOQYAFQDKNHLVMEYOPGG 180

Qy 181 DLSLLNRYEDQDENLIQFYLAELIIVAVSHVLMGVVHRDIKPNILVDRTHIKLVDF 240
Db 181 DFLSLLNRYEDQDENLIQFYLAELIIVAVSHVLMGVVHRDIKPNILVDRTHIKLVDF 240

Qy 241 GSAAKMNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKGTGLDCDMSVGVIAIYEMIYGR 300
Db 241 GSAAKMNSNK-VDAKLPIGTPDYMAPEVLTVMNDRRGTYGLDCDMSVGVIAIYEMVYK 299

Qy 301 SPFAEGTSARTFNINMFQRLKFPDPPKYSDFLDLIQSLLCQKRLKFEGLCCHPFF 360
Db 301 TPFTEGTSARTFNINMFQRLKFPDPPKYSSELDDLQSLLCVQKRLKFEGLCCHPFF 359

Qy 361 SKIDWNIRNSPPFVPTLSDDDTSNFDPEKNSWVSSSPCQLSPSGFSGEELPFVGF 420
Db 361 ARTDWNIRNSPPFVPTLSDDDTSNFDPEKNSWAFILCVPAEPLAFSGEELPFVGF 419

Qy 421 YSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQSDQKCHM 468
Db 421 YSKALGILGRSESVVSSLDSPAKYSMEKKLLIKSKELQSDQKCHV 467

RESULT 4
US-10-238-709-4
; Sequence 4, Application US/10238709
; Patent No. 6680188
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Mus Musculus
US-10-238-709-4

Query Match 20.6%; Score 2165.5; DB 4; Length 494;
Best Local Similarity 87.8%; Pred. No. 3.6e-104;
Matches 411; Conservative 26; Mismatches 30; Indels 1; Gaps 1;

Qy 1 MLKFYKGYARNPLDAGAAEPIASRASRLNLPFGKPPMTQQQMSPLSREGILDALFVLFE 60
```

Db 1 MLKFKYGVRRPPASASEPTIASASRLNLFQGGKPLMTQQQMSALSREGMLDALFALFE 60  
QY 61 ECQOPALMKIKHVSFNFRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVREKATG 120  
Db 61 ECQOPALMKIKHVSFNFRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVREKATG 120  
QY 121 DIYAMKWKKALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEYOPGG 180  
Db 121 DIYAMKWKKALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEYOPGG 180  
QY 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIPENILVDRTHIKLVDF 240  
Db 181 DFLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIPENILVDRTHIKLVDF 240  
QY 241 GSAAKXNSNMVNAKPIGTPDYMAPEVLTVMNGDGKTYGLDCDWSVGVYAYEVIYGR 300  
Db 241 GSAAKXNSNK-VDAKLPIGTPDYMAPEVLTVMNGDRRGTYGLDCDWSVGVYAYEVIYGR 299  
QY 301 SPEAGTARTFNINMFORFLKFPDDPKVSSDFLDLIQSLICGGKRLKFGLCCHPPF 360  
Db 300 TPTEGTSARTFNINMFORFLKFPDDPKVSSDFLDLIQSLICGGKRLKFGLCCHPPF 359  
QY 361 SKIDMNNIRNSPPFPVPTLKSDDTSNDFPEKNSWSSPCQLSPGSGBELPFGVGS 420  
Db 360 ARTDWNIRNSPPFPVPTLKSDDTSNDFPEKNSWAFILCVPAFLAFSGBELPFGVGS 419  
QY 421 YSKALGILGRSSVSGLSDPKTSMEKKLIKKELODSQDKCHKM 468  
Db 420 YSKALGILGRSSVSGLSDPKTSMEKKLIKKELODSQDKCHKV 467

RESULT 5  
US-09-916-204-2  
; Sequence 2, Application US/09916204  
; Patent No. 6638745  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEROF  
; FILE REFERENCE: CLO01184CIP  
; CURRENT APPLICATION NUMBER: US/09/916,204  
; CURRENT FILING DATE: 2001-07-24  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 257  
; TYPE: PRT  
; ORGANISM: Human  
US-09-916-204-2

Query Match 12.2%; Score 1277; DB 4; Length 257;  
Best Local Similarity 99.2%; Pred. No. 9.3e-59;  
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDAGAAEPIASASRLNLFQGGKPLMTQQQMSPLSREGILDALFVLFE 60  
Db 1 MLKFKYGARNPLDAGAAEPIASASRLNLFQGGKPLMTQQQMSPLSREGILDALFVLFE 60  
QY 61 ECQOPALMKIKHVSFNFRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVREKATG 120  
Db 61 ECQOPALMKIKHVSFNFRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVREKATG 120  
QY 121 DIYAMKWKKALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEYOPGG 180  
Db 121 DIYAMKWKKALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEYOPGG 180  
QY 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIPENILVDRTHIKLVDF 240  
Db 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIPENILVDRTHIKLVDF 240  
QY 241 GSAAKXNSNMVNAKPIGTPDYMAPEVLTVMNGDGKTYGLDCDWSVGVYAYEVIYGR 300

Db 241 GSAAKXNSNMVNAKPIGTPDYMAPEVLTVMNGDGKTYGLDCDWSVGVYAYEVIYGR 300  
QY 301 SPEAGTARTFNINMFORFLKFPDDPKVSSDFLDLIQSLICGGKRLKFGLCCHPPF 360  
Db 300 TPTEGTSARTFNINMFORFLKFPDDPKVSSDFLDLIQSLICGGKRLKFGLCCHPPF 359  
QY 361 SKIDMNNIRNSPPFPVPTLKSDDTSNDFPEKNSWSSPCQLSPGSGBELPFGVGS 420  
Db 360 ARTDWNIRNSPPFPVPTLKSDDTSNDFPEKNSWAFILCVPAFLAFSGBELPFGVGS 419  
QY 421 YSKALGILGRSSVSGLSDPKTSMEKKLIKKELODSQDKCHKM 468  
Db 420 YSKALGILGRSSVSGLSDPKTSMEKKLIKKELODSQDKCHKV 467

RESULT 6  
US-08-685-576-1  
; Sequence 1, Application US/08685576  
; Patent No. 5906819  
; GENERAL INFORMATION:  
; APPLICANT: Kaibuchi, Kozo  
; APPLICANT: Iwamatsu, Akihiro  
; APPLICANT: Nakano, Takeshi  
; APPLICANT: Ito, Masaaki  
; APPLICANT: Takahashi, No. 5906819uaki  
; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,576  
; FILING DATE: 24-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-325129  
; FILING DATE: 20-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-17150  
; FILING DATE: 05-JAN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-131206  
; FILING DATE: 26-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 16887/843  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1388 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-685-576-1

Query Match 11.9%; Score 1253; DB 2; Length 1388;  
Best Local Similarity 26.1%; Pred. No. 9.8e-57;  
Matches 394; Conservative 291; Mismatches 507; Indels 320; Gaps 50;

QY 15 GAAEPIA---SRASRLNLFQGGKPLMTQQQMSPLSREGILDALFVLFECSOPALMKIK 71  
Db 12 GAPEAVSGDGAGASR-----QRKLEALIRDRSPINVESLDDGLNPLVLDLDFPALRKNK 66  
QY 72 HVSNFVRKYSYDTIAELQELQPSAKDFEVRSLVGCCHFAEVOVREKATGDIYAMKWKKK 131  
Db 67 NIDNFLNRYEIKYKIRGLQMKAEYDVVKVIGRGAFGEVQLVVRHKASQKVYAMKLLSKF 126  
QY 132 ALLAQEQVFFFEERNILSRSTSPWIPOLQYAFQDKNHLVMEYOPGGDILSLNRYED 191  
Db 127 EMKRSDSAFFWEERDIMAFAFNSPWVQLFCAFDQDKYLYMVEYMPGGDLVNLMSNYD- 185



Db 23 SEVNSCLLDGLDLYDLDLPPALRKKNKIDNIDNLSRYKDTINKIRDLRMKABDEYEVVVI 82  
Qy 104 GCHFAEVQVREKATGDIYAMKMKKALLAQEQVSEFEERNILSRSTSPWLOQYA 163  
Db 83 GRGAFGEVQLVRHKSTKRVYAMKLLSKFEMIKESDSGAFFWEERDIPAFNSPWVQLFYA 142  
Qy 164 FQKKNHLYMEVYQPGGDLNLRNRYEDQDENLQFYLAELILAVSHVHMGYVHRDIK 223  
Db 143 FQDRYLYMWEYMPGGDLNLSNRYD--VPEKWARFYTAEVVLAIDAIHSGFTHRDYK 200  
Qy 224 PENILVDRTHIKILVDFGSAAKNSKMYNAKLPICTPDYMAPEVLTVMNGDGKGTGYGLD 283  
Db 201 PDNMLDKSGHLKADFGICMKMKGVRCVDTAVGTPTYISPEVLKSQGD--GYGRE 258  
Qy 284 CDWWSGVVYAYEMIGRSPFAEGBTSARTNNINMFORFLKFPDDPKVSSDFDLILQSLIC 343  
Db 259 CDWWSGVVLYEMLVGTTFYADSLVGTYSKMNHNKSLTFDDNDISKEAKNLLICAFLT 318  
Qy 344 GOKERLKFEF--LCCHPFESKID--WNNIRNSPPFPVPTLSDDDTSNFDPEKNWSVS 398  
Db 319 DREVLRNGVBEIKRHLFFKNDONAWETLRTVAVVPDLSDSDTSNFDLEEDKGE 378  
Qy 399 SS-PCOLSPSGSGEELPFGVSGYSKALGILGRSSVWSGLDSPAKTSSMEKXLLIKSE 457  
Db 379 ETFFPI---PKAFVGNQLPFVGTY-----YGNRRYLSSANPDNRNRTSSN-----ADKS 423  
Qy 458 LODS-ODKCKMEQEMTLRHRVSEVAVLQKEVELKASETOR---SILLEQD-LATVIT 512  
Db 424 LQESLOKTIYKUEEQ-----HNEMQLKDMEQCKRTSNIKLIDKIMKELD 468  
Qy 513 ECSSILRSLEQARMEYSQBDKALQLLHDIREQSRKLOEIKQEYQAOVE-EMRLMMNQL 571  
Db 469 EECGNORNLST---VSQIEKEKMLLQHRINEYQKAEQ--ENEXRNVENEVSTLKQOL 523  
Qy 572 EEDLYSARRSDIYSELSRESLAEEFKKATEQHKLLKAKDQCKPVEGYAKLEKIN 631  
Db 524 -EDLKVSQNSQ-----ANE---KLSQLQKLEANDLRTSTEDTAVLRKSH 568  
Qy 632 AQQLKIQLQKLEKAAERAELEKLQNRDSESGIRKVLVBAEERRHSLKNVKRL 691  
Db 569 TEMSKSISQL--ESLNLRELQER--NRILENSKSTQDXYQLQAILAEARRDRG-----618  
Qy 692 ETWERENLKDDIOTK---SQOIQOQADKILEEKEHREAOVSAQHLVHLKQEQHY 747  
Db 619 -----HDSMIGDLOARITSLQEEVKHLKHNLEKVEGERKEAQDMLNHE---KEK---666  
Qy 748 BEKIVLDNQIKDLADKETLENMMORHEEAEH---KGKILSOKAMINAMDSKIRSL 803  
Db 667 -----NNLEIDLNYK--LKSLOQLEQVEHNEHKVTARLTDKHQSIIEAKSVAMCEM 716  
Qy 804 EQRIVELSEANKLAANSLFTQRMKAQEMIS--ELROQKFYLETQAGKLEAQRKLEBQ 862  
Db 717 EKKLEEREAREKAENRVV---QIEKQCSMLDVLKQSQKLEHLTGKNE---RMEDE 768  
Qy 863 LEKISHQDHSKDNRLLELTRLEVSLEHBEQ-----LELKPQLTELQLS- 908  
Db 769 VKNLTQLQESQKRLLLQNLKTOAFADNKLGLEKQMKQBEINTLLLEAKK-LLEFELAQ 827  
Qy 909 -----LOERESOLTAQARAALLESOLROAKTELETTAAEAEETIQAHTADE 957  
Db 828 LTKQYRGNEGQRELQDLQLEABQYFSTLYKTQVKELKEIEKQENILKQLEQNEKET 887  
Qy 958 IQRKFDALRNSCTVITDLEEQNLQNTENAEILNNQNFYLSKOLDBAGANDEIVQLRSEV 1017  
Db 888 LATQLD-----LAETKAESQELA--RGLLEQYFELTQESKKAASRN-----927  
Qy 1018 DHLREITEREMQLTSQKQTEALATCTTMLBEQVMDLEALNDELLEKERQEWARSVLG 1077  
Db 928 ---RQEIFDKO-----HTVSRLEANSMLTKDLEILURNEELTEKOKKAEYKLEK 977  
Qy 1078 DE-----KSQFECRVRELQRLDTEKQSRADQRITESQVVELAVKHAELALQQA 1132  
Db 978 EEEISNLKAAFE-----KNINTERTLKTQAVNKLAEIMNRKQFKIDRKKANTQDLRKK 1030

Qy 1133 LKEQK---LKAESLSDKINDLEKHAMLEMNARSIOQKLETERELKORLLEBEQA---KLQ 1186  
Db 1031 EKENKQLQLELNQEREXFNQVVKH-----QKEL---NDMQAQLVEECAHRELQ 1077  
Qy 1187 QOMDLQKNHIFRLTCGLOEALDRADLLKTERSDLEYOLENIQVLYSHEKVMEGTISQOT 1246  
Db 1078 MQLASKESDIBQLRAKLLDLSSTSVASFPSAD-----ETDGNLPESR 1120  
Qy 1247 KLIDFLOAKMDQPAKKKKGLPSRRKEDPALPTQVPOVLOYNELKLALEKEKARCALEBALQ 1306  
Db 1121 -----IEGWSLVFNRGNIKRYGKKQYVVSSKILFYNDEQKEQNSVLDIDKL-- 1173  
Qy 1307 KTRIELSAREEAHRKATDHPHSTPATARQQIAMSALVRSRPHQPSAMSLAPPSSRR 1366  
Db 1174 ---FHVPRVTCQDYRAET-----EIEPKIFQILYANEDEC 1206  
Qy 1367 KESSTPEEFSRRLKERMVHNIPIHFVGLNMRATKCAVCLDTVHFGQASKCLEQVWCH 1426  
Db 1207 RQVEMEPVQAEKTNFNKHKGHEFIPTLYHFPANCDAKPLMHVFPKPPALECR-RCH 1265  
Qy 1427 PKCS-----TCLPATCGLPAEYATHFTFAFCRDKMNSPGLQTKPSSSLHLEGWM 1476  
Db 1266 VKCHRDHLDKKEDLIC-----PCKVSYDVTSA--RDMELLACSDQDEKQKVVHLV--K 1314  
Qy 1477 KVPNNKRG 1485  
Db 1315 KIPKNPFSG 1323

## RESULT 8

US-09-976-594-296  
; Sequence 296, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 296  
; LENGTH: 1388  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1  
US-09-976-594-296

Query Match 11.9%; Score 1249.5; DB 4; Length 1388;  
Best Local Similarity 26.3%; Pred. No. 1.5e-56;  
Matches 391; Conservative 291; Mismatches 513; Indels 293; Gaps 47;

Qy 24 ASRLNLFQCKPFFMTQQQSPLSRSGILLDALFVLEECSPALMKIKHVSFVXSDT 83  
Db 24 ASR-----QKRLALIRDPSPINVSLLDGLNSLVLDLDFPALRXNKNINDFLNRYEKI 78  
Qy 84 IAELOELQPSAKDFEVRSLVCGGHFAEVQVVRKATGDIYAMKMKKALLAQEQVSFFE 143  
Db 79 VKKIRGLQKAEEDYDVVKVIGRGAFGEVQLVRHKASQKYVAMKLLSKFEMIKESDSAFFW 138  
Qy 144 EERNILSRSTSPWLOQYAFQDKNHYLMEEYQPGDGLLSLNRVEDQDENLQFYLA 203  
Db 139 EERDIAFAFANSPWVQLFYAFQDDRYLYMVEYMPGGDLNLSNRYD--VPEKWARFYTA 196  
Qy 204 ELILAVHSHVLMGYVHRDIKPENILVDRTHIKILVDFGSAAKNSKMYNAKLPIGTDPY 263  
Db 197 EVVVALDAHSMGLIHRDVKPDNMLDKGHLKLAIDFGTCMKMDEIGMCHDITAVGTDPY 256

QY 264 MAFVLTVMGDKGTGYGLDCDWMWSGVYAYEMIVGRSPFAGTSGARTENNIMNFQRLK 323  
 Db 257 ISPEVLKSGQGD--GFGYRECDWMWSGVFLYEMLVGDTFTFYADSLVGTYSKIMDHNSLC 314  
 QY 324 FPDDPKVSDFLDIQSLCQKXERLKFEG--LCCHPEF--SKIDMNNIRNSPPFVPT 378  
 Db 315 FPDAEISKHAKNLCIAFLTDREVLGRNGVEIRQHPFFKNDQWHDNIRETAAPVVE 374  
 QY 379 LKSDDDTSNPDEPKNSW--VSSPCQLSPGSGEELPFVGSYSKALGICRSSESVSG 437  
 Db 375 LSSDIDSSNFDIEDDKGVETPPI---PKAFVGNQLPFGFTYRENNLLS----- 423  
 QY 438 LDSPA--KTSMEKKLLIKSLEQDQDKCKNCKQEMTLRRVRSVEAVLSQKEVELKA 495  
 Db 424 -DSPSCRENDISQSR---KNEESQETQKLYTLBEHLSN-----EMQA 462  
 QY 496 SETQSLLEQDLATYITECSSLRSLEQARMEVSQE-----DDKALQJLHDIRQSRKL 549  
 Db 463 KEE-----LEQ-----KCKSVNTRLEKTAKELEBEITLRKSVESALRQ-----ERKAL 507  
 QY 550 QETKEQEQYQAVEMRLMWNQLEEDIVASARRRSDLYESLRESRLAABEFKATECOHK 609  
 Db 508 LQHNKAEYQKADEADKGNLENDVNSLKQJEDLKKKNQNSQISTE-----KYNQLOEQ 563  
 QY 610 LLKAKQDGKPEVGEYAKLEKINAEQQLKIQ-----ELQEK--LEKAAKERAERELEKL 660  
 Db 564 LDETALLRTESDTAARLRTQAESKQIQLESNNRDLQKNCLLETAKLKEFEFINL 623  
 QY 661 QNRDSESGIRKLVAEERHSLKNVKELEMTWERNRLKDDIQTKSQJQQQWADKIL 720  
 Db 624 QSALESE---RRDTHGSETINDLQRIQLESDLLKNGKILLAKVELEKRRQJQERTD-L 679  
 QY 721 ELBEKREAOVSAQ--HLEVLHQQKE-QHYEEIKVLD-NOIKKOL--ADKETLENMQR 774  
 Db 680 EKEKSNWEIDMTYQKVIQSLQSEBAEHKATKARLADKNKIVESIEEAKSEAMKEMKK 739  
 QY 775 HEBEAHEKGI---LSEQKAMINAMDSKIRSLRQIVELSEANKLAANSLSLTQRMKQA 831  
 Db 740 LLEERTLKQVENLLLEAEKRCSLPCDLQSQSKINEL-----LROK 782  
 QY 832 BEMISLROQKYLETOAGLEQNRKLEBQLEKSHQDHSKRNRLLETRLERLVSLRH 891  
 Db 783 DVNEDVRNUTLKIEQETQKCCLTQNDLQKQVNTLQKSEK-QLKQENHLMEMKXNL 841  
 QY 892 BEQKLEKRLQTELQSLQRESQLTALQARALESQRLQAKTELEETT---ABABEEI 948  
 Db 842 EKQNAELRKERQDAGQMKELQLEAEQVFTLYKTQVRELKEECEETKLKGLQKQK 901  
 QY 949 QALTAHDETORFEDALRNSCTVITLLEQNLQITEDNNAELNNQNFYLSKQLEASGAND 1008  
 Db 902 QELQDERDSLAAQLE-----ITLTKADSEQLARSIAEBEQVSDLEKIKME----- 947  
 QY 1009 EIVQLRSEVDHLREITEREMQLTQKQMEALTKTTCTMLEEQVMDLEALNDELLEKERQ 1068  
 Db 948 --LEIKEMWARHQELTEKDATIASLEETNRLTSDVANLAN--EKEELNNKLKDVQSQ 1002  
 QY 1069 WEAWRSVLGE-----KQOFECVRRELQRLDTEKQSPARADQRTESRQVVELAVKE 1121  
 Db 1003 ----LSRLKDEEISAAIAKAQFE-----KQLLTERTLTKTQAVNKL----- 1038  
 QY 1122 HKAEILAQALKEQKLKABSLDKLNDLEKHAMLEMMARSQOQKLETERELKQRLLEE 1181  
 Db 1039 --AEMWNEKEPVK-----RGNDTDVRKK--EKENRKLHMLKSERE---KLTQQ 1080  
 QY 1182 QAKLQQQMDLQKNHIFLQTO---GLQALDRALLKTERSLEYQLENQVLY----- 1231  
 Db 1081 MIXYQKELNEMQAIAESQIRIELQMTLDSKQ-----SDIEQLRSQALHIGLIDSSS 1134  
 QY 1232 -----SHEKVKMEGTTS-----QOTKLIDFLQAGWDQPAKKKGLFSRRKEDPA 1275  
 Db 1135 IGSFGDAEADDGFPESLELGSWLSLPVYNTTKFGWVKVYIVVSSKILFYDSEQKEQS 1194

QY 1276 LPTQV-----PQYNELKLALKEKARCACAELEELQKTRIELRSAREBAHRKAT 1325  
 Db 1195 NPYMVLDDIKLFHVRPVTQTDVYRADAKETPRIFQI----- 1230  
 QY 1326 DHPFSPATARQOIAMSAIVRSPHQPSAMSLAPPSSRRKESSTP-----BFFSRRLKE 1381  
 Db 1231 -----LYANEGSKKKEQEPFVFPVGEKSNYICH 1258  
 QY 1382 RMHNIHPRNFVGNMKAATCAVCLDTVHFGRQASKCQVNMCHPKC 1429  
 Db 1259 KGHEFIFTLHYF-----PTNCEACMKPLMHMPKPPPALECR-RCHIKC 1300

RESULT 9  
 US-08-685-576-4  
 ; Sequence 4, Application US/08685576  
 ; Patent No. 5906819  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kaibuchi, Kozo  
 ; APPLICANT: Iwamatsu, Akihiro  
 ; APPLICANT: Nakano, Takeshi  
 ; APPLICANT: Ito, Masaaki  
 ; APPLICANT: Takahashi, No. 5906819uaki  
 ; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 3000 K Street, N.W., Suite 500  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/685,576  
 ; FILING DATE: 24-JUL-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 7-325129  
 ; FILING DATE: 20-NOV-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 8-17150  
 ; FILING DATE: 05-JAN-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 8-131206  
 ; FILING DATE: 26-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bent, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 16887/843  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202)672-5300  
 ; TELEFAX: (202)672-5399  
 ; TELEX: 904136  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1388 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-685-576-4

Query Match 11.9%; Score 1247.5; DB 2; Length 1388;  
 Best Local Similarity 26.3%; Pred. No. 1.9e-56;  
 Matches 391; Conservative 290; Mismatches 54; Indels 293; Gaps 47;

QY 24 ASRLNLFQGGPPMTQQQMSPLSREGILDALFVLFECQSPALMKJHVNFRVKYSDT 83  
 Db 24 ASR-----QRKLEALIRDPSPINVESLLDGLNSLVLDLDFPALRKNKNDINFLNRYEKI 78



```

; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-916-204-5

Query Match      11.2%; Score 1170; DB 4; Length 251;
Best Local Similarity 89.6%; Pred. No. 2.9e-53;
Matches 225; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSRREGILDALFVLFE 60
DB 1 MLKFKYGVNRNPPASASEPIASRASRLNLFQGGKPLMTQQQMSALSREGMLDALFALFE 60

QY 61 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120
DB 61 ECSQPALMKHVSFVQKYSYDTIAELRELQPSARDFEVRSLVGCCHFAEVQVVRKATG 120

QY 121 DIYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLIYLMEEYQPGG 180
DB 121 DVYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLIYLMEEYQPGG 180

QY 181 DLISLNRVEDQDENLQFYLAELILAVHSVHMGYVHRDIKPENILVDRGTGHIKLVDF 240
DB 181 DFJSLNRYEDQDENLQFYLAELILAVHSVHMGYVHRDIKPENILVDRGTGHIKLVDF 240

QY 241 GSAAKMNSNKM 251
DB 241 GSAAKMNSKV 251

RESULT 12
US-09-916-204-5
; Sequence 6, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-916-204-6

Query Match      11.2%; Score 1170; DB 4; Length 251;
Best Local Similarity 89.6%; Pred. No. 2.9e-53;
Matches 225; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSRREGILDALFVLFE 60
DB 1 MLKFKYGVNRNPPASASEPIASRASRLNLFQGGKPLMTQQQMSALSREGMLDALFALFE 60

QY 61 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120
DB 61 ECSQPALMKHVSFVQKYSYDTIAELRELQPSARDFEVRSLVGCCHFAEVQVVRKATG 120

QY 121 DIYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLIYLMEEYQPGG 180
DB 121 DVYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLIYLMEEYQPGG 180

QY 181 DLISLNRVEDQDENLQFYLAELILAVHSVHMGYVHRDIKPENILVDRGTGHIKLVDF 240
DB 181 DFJSLNRYEDQDENLQFYLAELILAVHSVHMGYVHRDIKPENILVDRGTGHIKLVDF 240

QY 241 GSAAKMNSNKM 251
DB 241 GSAAKMNSKV 251

RESULT 13
US-08-630-822A-62
; Sequence 62, Application US/08630822A
; Patent No. 5840695
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDA
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
; TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,822A
; FILING DATE: 11-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CONNELL, GARY J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-630-822A-62

Query Match      9.9%; Score 1043.5; DB 2; Length 900;
Best Local Similarity 27.0%; Pred. No. 3.7e-46;
Matches 279; Conservative 221; Mismatches 364; Indels 171; Gaps 31;

QY 70 IKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATGDIYAMKVMK 129
DB 1 MKSEAYTNRTETITASEIVNLRMKPDDFNLIKVGAFGEVQLVRHKSTAQVFAVKRLS 60

QY 130 KALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLIYLMEEYQPGDLSILNRY 189
DB 61 KPEMIKRPDSAFFWEERHIMAHAKSEWIVQLHFAFQDKYLYMYMDYMPGGDLVSLMSDY 120

QY 190 EDQDENLQFYLAELILAVHSVHMGYVHRDIKPENILVDRGTGHIKLVDFGSAKNSN 249
DB 121 E--IPEKWMFYTMEVVLALDTIISMGFVHRDVFPDNMLDKYGLKLADEGTCKMDTD 178

QY 250 KMVNAKPIGTPDYMAPEVLTVNMGDKGTYGLDCDMMWSVGVIAYEMIVYGRSPFAEGTSA 309

```

[illegible]

Db 504 --KGELEHN---QKLLKQAVELR-----VAQSA-TEQLNNELQETWQGLQT 543  
Qy 644 KLEKAARERAELEKLNREDSSEIRKLVAEERHRSLENKVKRLTMRERENRDKD 703  
Db 544 Q-----RDALQOEVASLQKLSQERSRS---QASDMQIELEAKLQALHIELEHVRNCD 595  
Qy 704 DIQTSQOIQOMADKILEEKEHREAVSAQHLEVLHKKOEQHYEKKIKVLDNQIKDLA 763  
Db 596 KV---TQDNQRLERISTLEKE-----CASLELELKATQNKYEQEVK-----A 635  
Qy 764 DKETLENMORHEEAEHKGKILSEQKAMINAMDSKIRSLERIVELSEANKLAANSLF 823  
Db 636 HRET-----EKSRLVSKEEANN---EYKALQIKLNEEKSAQKSDQNS-- 676  
Qy 824 TQNNKAQEMISELROCKFYLETQAGKLEAQNKLKEOLEKISHQDHSKRNLELETR 883  
Db 677 -----QEK---ERQISMLSVDYRQIQOQLKLEGEYRQESKVALHSQ 717  
Qy 884 LREVSLEHEBQKLEKRLQTELQSLQERESQLTALQAAARAALESQLOAKTELETTAE 943  
Db 718 IEQQLKKSQLOSELGVQ-----RSQTAHLTAREAQVGEVAHLRDAKRNVEELHK 769  
Qy 944 ABEIQAHTARHDEIQKFDALNSCTVI-----TDLEEQLNQLTEDNAELNNQNFYLSQ 999  
Db 770 LKTARSDVNAQMKELQOEVAEQVFTLYKTHSNELKEELEKSRHIOEMEERESLVHQ 829  
Qy 1000 LDEASGANDEIVQLRSEVH-----LRREITERMOLTSQKQWMEALKTTCTMLEBQVMDL 1055  
Db 830 LQIALARADSEALARSIADESIADLEKXTMKELEL-----KEULNKNRTELSQKDISI 883  
Qy 1056 EALNDELLEKEROME 1070  
Db 884 SALRERENEQKLE 898

RESULT 15  
US-09-171-156A-21  
; Sequence 21, Application US/09171156A  
; Patent No. 6368846  
; GENERAL INFORMATION:  
; APPLICANT: Hunter, Shirley Wu  
; Weber, Eric R.  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND  
; APPARATUS TO COLLECT SUCH PROTEINS  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS P.C.  
; STREET: 1560 BROADWAY, SUITE 1200  
; CITY: DENVER  
; STATE: CO  
; COUNTRY: U.S.A.  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/171.156A  
; FILING DATE: 04-Mar-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cornell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 900 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-171-156A-21  
Query Match 9.9%; Score 1043.5; DB 4; Length 900;  
Best Local Similarity 27.08; Pred. No. 3.7e-46;  
Matches 279; Conservative 221; Mismatches 364; Indels 171; Gaps 31;  
Qy 70 IKVSNFVRKYSDTIALQELQPSAXDFEVSIVGVGCHFAEVGVREKATGDIYAKVMK 129  
Db 1 MKSEATYNYRIEIASIVNLRMKPDDFNLIKVGAGFGEVOLVHKSTAQVFAKRLS 60  
Qy 130 KKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDKNHLYLMEYQPGDLSLLNRY 189  
Db 61 KFEIMIKRPDSAFWEERHIMAHAKSEWIVQLHEAFQDKYLYVMVDMYMPGGLVLSMSDY 120  
Qy 190 EDQDENLIFYLAELILAVSHVHLGYVHRDIKPENILVDRTHIKLVDFGSAAKWNSN 249  
Db 121 E--IPEKAMFYTMVVYALDTHSMGFVHRDVKPNMLLDKYLHKLADFGTCKMMDID 178  
Qy 250 KMNAKLPIGTDPYMAPEVLTVMNGDKGTGYGLDCDWSVGVYAYEMIVGRSPFAGTSA 309  
Db 179 GLVRSNNAVGTPDISPEVL--SQSGEGYVGRECDWSVGIIFYEMLFGETPFYADSLV 236  
Qy 310 RTFNININFORFLKPPDDPKVSSDFLDLIQSLICGQKERL---KFEGLCCHPFPSKIDW- 365  
Db 237 GTYSKIMDHNSLTFPPEVEISQYARSLIQGFLTDRTQRLGRNEVEIEKHHPFFINDQWT 296  
Qy 366 -NNIRNSPPPEVPTLKDDDTSNFDE-----PEKISWVSSPCQLSPSGSGEELPFVG 418  
Db 297 PDNLRDAPPVPFELSDDDTFRNFDIERDETPEENFPI-----PKTFAGNHLFPVG 348  
Qy 419 FSYSKALGILGRSESVVSGLDSPAKYSSMEKKLLIKSKE-----LQDSQDKCHMEQ 470  
Db 349 FTYNGDYQLLTN-----GGVRNSDMVDTKLNNICVSSKDDVLNQLNLLQEKGNSENKLT 403  
Qy 471 EMTLHRRVSEVAEVLQKEVELKASQTSRSLQEDLATYITCSSLKRSLE---QARME 527  
Db 404 NTQLLSNKLDE---LQRECELR---NQAGDYKEKLTFKLSCKELQKAEFENELRRK 456  
Qy 528 VSQ---BDDKALQLLHDIREQSKLQEI-KEQFYQAQVEEMRLMMQLEEDLVSAARRSD 583  
Db 457 TESLLVETKKRLD---EEQNKRTREMNNOQHNDKINMLEKQINDLQEKL----- 503  
Qy 584 LYSELSRESRLAAEFKFKATEQHKLLKAKQDQKPEVGEYAKLEKINAEQKQIKLOE 643  
Db 504 --KGELEHN---QKLLKQAVELR-----VAQSA-TEQLNNELQETWQGLQT 543  
Qy 644 KLEKAARERAELEKLNREDSSEIRKLVAEERHRSLENKVKRLTMRERENRDKD 703  
Db 544 Q-----RDALQOEVASLQKLSQERSRS---QASDMQIELEAKLQALHIELEHVRNCD 595  
Qy 704 DIQTSQOIQOMADKILEEKEHREAVSAQHLEVLHKKOEQHYEKKIKVLDNQIKDLA 763  
Db 596 KV---TQDNQRLERISTLEKE-----CASLELELKATQNKYEQEVK-----A 635  
Qy 764 DKETLENMORHEEAEHKGKILSEQKAMINAMDSKIRSLERIVELSEANKLAANSLF 823  
Db 636 HRET-----EKSRLVSKEEANN---EYKALQIKLNEEKSAQKSDQNS-- 676  
Qy 824 TQNNKAQEMISELROCKFYLETQAGKLEAQNKLKEOLEKISHQDHSKRNLELETR 883  
Db 677 -----QEK---ERQISMLSVDYRQIQOQLKLEGEYRQESKVALHSQ 717  
Qy 884 LREVSLEHEBQKLEKRLQTELQSLQERESQLTALQAAARAALESQLOAKTELETTAE 943  
Db 718 IEQQLKKSQLOSELGVQ-----RSQTAHLTAREAQVGEVAHLRDAKRNVEELHK 769  
Qy 944 ABEIQAHTARHDEIQKFDALNSCTVI-----TDLEEQLNQLTEDNAELNNQNFYLSQ 999  
Db 770 LKTARSDVNAQMKELQOEVAEQVFTLYKTHSNELKEELEKSRHIOEMEERESLVHQ 829

QY 1000 LDEASGANDEIVQLRSEVDH---LRRREITEREMOLTSQKQTIMEALKTTCTMLEEQVMDL 1055  
Db 830 LQIALARADSEALARSIADESIADLEKXKTWKELEL-----KELLNKNRTELSQKDISI 883  
QY 1056 EALNDELLEKERQWE 1070  
Db 884 SALRERENEQKKLE 898

Search completed: July 3, 2004, 10:18:31  
Job time : 29 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:07:35 ; Search time 36 Seconds  
(without alignments)  
5485.589 Million cell updates/sec

Title: US-10-017-216-2  
Perfect score: 10490  
Sequence: 1 MLKFKYGARNPLDAGAEPI.....QLNGEIRQQVEKSVLRDYC 2053

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7491.5	71.4	1597	2 S68420	citron - mouse
2	1450	13.8	1732	2 T14039	protein kinase (EC
3	1372	13.1	1702	2 T14050	protein kinase (EC
4	1261.5	12.0	1354	2 S74244	serine/threonine-s
5	1253	11.9	1388	2 S70233	serine/threonine-s
6	1250.5	11.9	1354	2 S69211	serine/threonine-s
7	1245	11.9	1388	2 S74245	serine/threonine-s
8	1245	11.9	1548	2 T25808	hypothetical prote
9	1112.5	10.6	1173	2 T25539	hypothetical prote
10	878.5	8.4	624	2 B49364	protein kinase (EC
11	862	8.2	557	2 S71829	serine/threonine-s
12	739.5	7.0	1356	2 T16718	hypothetical prote
13	712	6.8	522	2 G86431	protein kinase T5I
14	706.5	6.7	479	2 S42864	protein kinase (EC
15	706.5	6.7	1256	2 T26101	hypothetical prote
16	700	6.7	441	2 T78393	myotonic dystrophy
17	700	6.7	474	2 T78396	myotonic dystrophy
18	700	6.7	516	2 T78394	myotonic dystrophy
19	698	6.7	756	2 S60966	probable protein k
20	691	6.6	469	2 T41723	serine/threonine-s
21	691	6.6	526	2 S49077	protein kinase PKR
22	687	6.5	569	2 A86170	hypothetical prote
23	687	6.5	596	2 F84589	probable protein k
24	683.5	6.5	1286	2 T16507	hypothetical prote
25	667.5	6.4	443	2 D71405	probable protein k
26	667.5	6.4	475	2 H85156	protein kinase (im
27	665.5	6.3	465	2 I38133	protein kinase (EC
28	661	6.3	500	2 S42867	protein kinase (EC
29	651.5	6.2	483	2 T05188	protein kinase F4I

30	630.5	6.0	665	2 S70706	probable protein k
31	630	6.0	480	2 T47255	serine/threonine k
32	630	6.0	545	2 T01288	protein kinase F27
33	630	6.0	598	2 T47254	serine/threonine k
34	630	6.0	620	2 S22711	probable protein k
35	621.5	5.9	412	2 T78395	myotonic dystrophy
36	599	5.7	908	2 T25035	hypothetical prote
37	594	5.7	1039	2 A56155	tumor suppressor p
38	582.5	5.6	1410	1 A57013	early endosome ant
39	573.5	5.5	2139	2 T18296	myosin heavy chain
40	570	5.4	1909	2 A45592	liver stage anti
41	565	5.4	1738	2 T14867	interaptin - silme
42	559	5.3	2442	2 T08621	centrosome associa
43	541.5	5.2	624	2 T41341	probable serine-th
44	540.5	5.2	2116	2 A26655	myosin heavy chain
45	540	5.1	607	2 S62556	probable serine/th

ALIGNMENTS

RESULT 1  
S68420  
citron - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 05-Nov-1999  
C:Accession: S68420  
R:Madaule, P.; Furuhashiki, T.; Reid, T.; Ishizaki, T.; Watanabe, G.; Morii, N.; Narumiya  
FEBS Lett, 377, 243-248, 1995  
A>Title: A novel partner for the GTP-bound forms of rho and rac.  
A:Reference number: S68420; MUID:96128238; PMID:8543060  
A:Accession: S68420  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1597 <MAD>  
A:Cross-references: EMBL:U39904; NID:G1079733; PIDN:AAC52341.1; PID:G1079734  
C:Superfamily: protein kinase C zinc-binding repeat homology; pleckstrin repeat homology  
C:Keywords: alternative splicing  
F:931-979/Domain: protein kinase C zinc-binding repeat homology <K22>

Query Match 71.4%; Score 7491.5; DB 2; Length 1597;  
Best Local Similarity 93.1%; Pred. No. 2.3e-199;  
Matches 1492; Conservative 16; Mismatches 12; Indels 83; Gaps 4;

Qy	468	MEQMTLHRRVSEVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSIKRSLEQARME	527
Db	9	MEQMTLHRRVSEVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSIKRSLEQARME	68
Qy	528	VSQEDDKALQLLHDIREQSRKLOEIKEOEQAOVEENRLMMNQLLEEDLVGARRSRDLYES	587
Db	69	VSQEDDKALQLLHDIREQSRKLOEIKEOEQAOVEENRLMMNQLLEEDLVGARRSRDLYES	128
Qy	588	ELRESRLAAEFKRRKATECOHKLKAKDQKPEVGEYAKLEKINAFQOLKIQLQEKLEK	647
Db	129	ELRESRLAAEFKRRKANECQKLMKAKDQKPEVGEYSKLEKINAEQOLKIQLQEKLEK	188
Qy	648	A-----AKERAEREKLNRRDSSGIRKKLVEAEERHSLENKVRL	691
Db	189	AVKASTATELLQMIROKAREREKLNRRDSSGIRKKLVEAEERHSLENKVRL	248
Qy	692	ETMERRENRLKDDIQTQKSQIQQMADKILELEKHEAQAQVSAQHLEVLKQKEQHYBEKI	751
Db	249	ETMERRENRLKDDIQTQKSQIQQMADKILELEKHEAQAQVSAQHLEVLKQKEQHYBEKI	308
Qy	752	KVLNQIKKLIADKETLENMQRHEEBAHEKGLTSQKAMINAMDSKIRSLEQRIVELS	811
Db	309	KVLNQIKKLIADKESLENMQRHEEBAHEKGLTSQKAMINAMDSKIRSLEQRIVELS	368
Qy	812	EANKLAANSLSFTORNKAQBEIMSELROQKFYLETQAGKLEAQRNKLQLEKIKISHQDH	871
Db	369	EANKLAANSLSFTORNKAQBEIMSELROQKFYLETQAGKLEAQRNKLQLEKIKISHQDH	428
Qy	872	SDKNRLLETRLEVRVSLSEHEEQKLEKLKQLITELQLSLQERESQLTALQAARAALESQLR	931

Db 429 SDKRLLETRREVSLEHEOKLEKRLQTLQSLQRESQTLAQARALESQUR 488  
QY 932 QAKTELETTAAABEEIQAHTAHRDEIQKFDALRNSCTVITDLERQNLQTLTEDNAELNN 991  
Db 489 QAKTELETTAAABEEIQAHTAHRDEIQKFDALRNSCTVITDLERQNLQTLTEDNAELNN 548  
QY 992 QNFVLSQOLDEASGANDIEIQLSEVDHLRREITEREMQLTSQKQTMWALKTCTMLBQ 1051  
Db 549 QNFVLSQOLDEASGANDIEIQLSEVDHLRREITEREMQLTSQKQTMWALKTCTMLBQ 608  
QY 1052 VMLEALNDLLEKKEQWAEWSVLGDEKSOFCRVREIQRLMDLTKQGRADADQRTES 1111  
Db 609 VMLEALNDLLEKKEQWAEWSVLGDEKSOFCRVREIQRLMDLTKQGRADADQRTES 668  
QY 1112 QOVVELAVKHKAEIILALQALKEQIKAEBSLSDKLNLEKXHAMLENNARSLOQKLETE 1171  
Db 669 QOVVELAVKHKAEIILALQALKEQIKAEBSLSDKLNLEKXHAMLENNARSLOQKLETE 728  
QY 1172 RELKORLLEBQAKLOQMDLOKXHI FRLTQGLQALDRADLLKTERSLEYOLENIQVLY 1231  
Db 729 RELKORLLEBQAKLOQMDLOKXHI FRLTQGLQALDRADLLKTERSLEYOLENIQVLY 788  
QY 1232 SHEKVMKEGTISOOTKLIDFLQAMQOPAKKKGLFSRRKEDPALPTQVPLQYNELKAL 1291  
Db 789 SHEKVMKEGTISOOTKLIDFLQAMQOPAKKK-----VPLQYNELKAL 833  
QY 1292 EKKARCAELEALOKTRIELRSAREEAAHRAKADHPHSTPATARQOIAMSALVRSPEH 1351  
Db 834 EKKARCAELEALOKTRIELRSAREEAAHRAKADHPHSTPATARQOIAMSALVRSPEH 893  
QY 1352 QPSAMILLAPPSRRKESSTPEFSRRLKERHNNIPHRFNVLGNMRAKCAVCLDTVHF 1411  
Db 894 QPSAMILLAPPSRRKESSTPEFSRRLKERHNNIPHRFNVLGNMRAKCAVCLDTVHF 953  
QY 1412 GRQASKLEQVWCHPKCSTCLPATCGLPAEYATHTEAFCDKKNPSGLOQKPESSSLH 1471  
Db 954 GRQASKLEQVWCHPKCSTCLPATCGLPAEYATHTEAFCDKKNPSGLOQKPESSSLH 1013  
QY 1472 LEGWMKVPNNKGGQGWDRKYIVLGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIH 1531  
Db 1014 LEGWMKVPNNKGGQGWDRKYIVLGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIH 1073  
QY 1532 GAVGASELANTAKA----- 1545  
Db 1074 GAVGASELANTAKADVPYILKMHSPHTTWCPTGLYLLAPSPDKQRWTALESWAGG 1133  
QY 1546 ---EKAEADAKLLGNSLLKLEGGDRDLNMNCTLPFSDDQVVLVGTBEGLYALNVLKNSLTH 1601  
Db 1134 RVSREKAEADAKLLGNSLLKLEGGDRDLNMNCTLPFSDDQVVLVGTBEGLYALNVLKNSLTH 1193  
QY 1602 VPIGAVFOIYIKOLEKLLMIAGBERALCLVDVKVKQSLAQSHLPAQPDSPNIFEAV 1661  
Db 1194 VPIGAVFOIYIKOLEKLLMIAGBERALCLVDVKVKQSLAQSHLPAQPDSPNIFEAV 1253  
QY 1662 KCHFLFAGKINGLCICAAMPSKVILRYNENLSKYCIKKEIETSEPCSCIFHFNYSIL 1721  
Db 1254 KCHFLFAGKINGLCICAAMPSKVILRYNENLSKYCIKKEIETSEPCSCIFHFNYSIL 1313  
QY 1722 IGTNKFYEIDMKQYTLDEFLDNDSLAPAVPAASNSFPFVSIVQVNSAGOREEYLLCFH 1781  
Db 1314 IGTNKFYEIDMKQYTLDEFLDNDSLAPAVPAASNSFPFVSIVQVNSAGOREEYLLCFH 1373  
QY 1782 EREGVDSVGRSRTDILKWSLPLAFAYREPLYFVTHFNLSLEVIEIQARSAGTPARAY 1841  
Db 1374 EREGVDSVGRSRTDILKWSLPLAFAYREPLYFVTHFNLSLEVIEIQARSAGTPARAY 1433  
QY 1842 LDIPNRYLGPALISSGAIYLAASSYQDKLRVICCKGNLVKESGTEHRRGPSTRSSPNKRG 1901  
Db 1434 LDIPNRYLGPALISSGAIYLAASSYQDKLRVICCKGNLVKESGTEHRRGPSTRSSPNKRG 1493  
QY 1902 PPTYNHITKRVASSPAPPEGSHRPRESTPHRY--REGRTELRDKSPGRLERKSPG 1959

Db 1494 PPTYNHITKRVASSPAPPEGSHRPRESTPHRYDRGRTELRRDKSPGRLERKSPG 1553  
QY 1960 RMLSTRERSGRLEFEDSSRGRLPAGAVRTPLSQVNKRGQSA 2002  
Db 1554 RMLSTRERSGRLEFEDSSRGRLPAGAVRTPLSQVNKVDQSS 1596  
RESULT 2  
T14039  
protein kinase (EC 2.7.1.37), myotonic dystrophy-associated - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jun-2000  
C:Accession: T14039  
R:Jeung, T.; Chen, X.Q.; Tan, I.; Manser, E.; Lim, L.  
Mol. Cell. Biol. 18, 130-140, 1998  
A:Title: Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector  
A:Reference number: Z17862; MUID:98078670; PMID:9418861  
A:Accession: T14039  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1732 <LEU>  
A:Cross-references: EMBL:AF021935; NID:92736150; PID:92736151; PIDN:AAC02941.1  
C:Genetics:  
A:Gene: MRCK  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; phosphotransferase  
F:75-343/Domain: protein kinase homology <KIN>  
F:1013-1062/Domain: protein kinase C zinc-binding repeat homology <KZN>  
Query Match 13.8%; Score 1450; DB 2; Length 1732;  
Best Local Similarity 24.6%; Pred.No.3.4e-33;  
Matches 521; Conservative 319; Mismatches 648; Indels 634; Gaps 75;  
QY 36 PFMTQQQMSPLSREGILDALFLVFEECSQPALMKIKHVSNFVKYSDDTIAELQELQPSAK 95  
Db 18 PAQTNGQC--FSVETLLDILICLYDECNPSPLREKNILLEYLWAKPFTSKVKQMLHRE 75  
QY 96 DFEVRSVLGGCHFAEVQVREKATGIYAMKVMKKALLAQEQVSFFEEBENILSRSTSP 155  
Db 76 DFEILKVIKRGAFGEVAVVVKLNADKVFAMKILNKWMLKRAETACFBEERDVLVNGDSK 135  
QY 156 WIPQLQYAFQKQKHLIMEEYQPGDLLSLINRYEQLDENLIQFYLAELILAVHSVLM 215  
Db 136 WITTLHAFQDDNNLYLWMDYVYGGDLLTLLSKFEDELPEEMARFYLAEMVIAIDSVHL 195  
QY 216 GYVHRDIKPNILVDRTGHIKLVDFGSAKNGNKNMNAKLPIGTDPDYVAPEVLTWNGD 275  
Db 196 HYVHRDIKPNILMDXNGHILRADFGSKLMDGTQVSSVAVGTPTYISPEILQAVE-D 254  
QY 276 GKTYGLDCCWNSGVYAYEMIVYGRSPFAGTGAFTFNIMNFORFLKPPDD--PKYSSDF 334  
Db 255 GKRGYCECDWNSGLVCMYEMLYGETPFYAESIVETYGKIMNHERFQFPQTQVTDVSENA 314  
QY 335 LDLIQSLCCOKERLKEFGL---CCHPFTSKIDWNIRNSPPPEVPTLKSDDDDTSNDEP 391  
Db 315 KDILIRLICSREHRLQNGIEDFKHPFGSGIDWNIRNCEAPIYIEVSSPTDTSNFDVD 374  
QY 392 E---KNSVSSSPQCLSPSGFGSEELPFVGFYSKALGILGRS-----ESV 434  
Db 375 DDCLKNSETMPPF--THTAFSGHLLPFVGFVTSVCLSDRSCLRTVAGT\*SLDLNVN 431  
QY 435 VSGLDSPAKTSMEKLLIKSKELQDSQDKCHMEQEMTLHRRVSE---VEA----- 484  
Db 432 QRTLDNNLATEATERRI-----KKLEQEKLELTKLOESTQTQVALQYSTV 477  
QY 485 ---VLSQKEVELKASRTQSLLEQDLATYITCSSLKRSLEQARMEVSGEDDKALQLLHD 541  
Db 478 DGPLTASKDLKLEIKSLETEKLRKQVA---EVNHLQOOLEAN--SVRRELDAAFRQIA 532  
QY 542 IRECSRLQEIKEQEQVQAQVEEMRLMMNOLEDLVASRRSDLYESELESRLAABEFKR 601  
Db 533 FEKQIKTLQERE-----ENKELVQASERLKNQSKELKDA----- 568



Db 203 PDNVLLDVNGHRLADFGSCLXNDDGTQSSVAVGTPDYISPEILQAME-DGMSKYGPE 261  
Qy 284 CDWWSGVYIAYEYGRSPFAGTSARTNNINWFQFLKFPDD-PKYSSPFLDIQSL 342  
Db 262 CDWWSLGVCMYEMLYGETPFYAESLVETYGKIMNHEERFPFPHSHVTDVSEAKDILQRLI 321  
Qy 343 CGQKERLKEGGL--CCHPFFSKIDANNIRNSPPFPVPTLKSDDDTSPFDEP---KNSW 396  
Db 322 CGERERLQNGHIEDFKKHAFFEGELNENIRNLEAYIPDVVSPSTSDTFDVEDDVLRIE 381  
Qy 397 VSSSPQQLSP--SGFSGBELPFVGSYSKALGILGRSESVVGLDSPAQTSSMEKLLI 453  
Db 382 I-----LPPGSHGTGSLHLFPFGTFT-----TESCF--DRGSLKSMIQSNTLT 425  
Qy 454 KSELODSQDKCHKMEQEMTRLHRRVSEVEAVLSQVELKASEQSRSLLEQDLATYITE 513  
Db 426 KDQEDVDRLNSLOIAYERRR----- 449  
Qy 514 CSSLKESLQARMEVS---QEDDKALQLLHDIRE---QSRKLEIKOEYQAOVEMRLM 567  
Db 450 -----LQEKLESLKLOESTQTQSLHGSTALGNSNRDKTEK--RLNEELERMKS 500  
Qy 568 M-----NOL-----EDLVARRSDLYSELSERSLAABEFKFKKATECQHKLLKAKDOGKPEV 621  
Db 501 MADSNLERQLEDTVTLRQ----- 519  
Qy 622 GEYAKLEKINAEOLKIQLQEKLEKAAKERAERLEKLNQREDSSGIRKKLVAEERR 681  
Db 520 -----EHDSTQRLKLEKQVRLAQEKEB-----LHKQVLEASER- 555  
Qy 682 HSLNKVKLEWNERENLKDDIQTKSOIQOMADKILEBEKREBAQVSAHLEVLK 741  
Db 556 -----LKSOTKELKAHQKRALQFSS-----ELNERMAELRSQKQVSRQLR 599  
Qy 742 QKEQHYEEKIKVLNDNOIKDLADKETLENMWRHBEAEKGIKILSEQKAMINAMDSKIR 801  
Db 600 DKEEEMEVAKIKDS--MRQIRKSE-----KSKLEAR----- 632  
Qy 802 SLQRIVELSEANKLAANSLFTQRMKAQEMISELROOKFYLETOAGKLEAQNKLBE 861  
Db 633 -LEDAAEASKERKLREHSEFS-----KQERELETQV--KQGRGPGCATLEHQ 681  
Qy 862 QLEKISHQSDKRNLELETRLSVLSHEEQKLELR-----QLTELQSLQERESQ 917  
Db 682 ETSKIRS-----ELEKKV-----LFYEBELVRRERSHVLEKVNKVEHSESHQ 727  
Qy 918 ALQAAARALESQURQAKTELEETTAEEBEOIALTAHRDEIQKFDALRNSCTVITDLE 977  
Db 728 ALQKEVLMKDKLEKSK--RERHSEVEBAIGAM--KQYER----- 764  
Qy 978 QLNQLEDNALNNQNYLSKQIDRAGAN-----DEIVQIRSE---VDHLREITEREMQ 1030  
Db 765 ERAMLFDENKLLTAENEKLCFSYDKLTAQNRQLEDELDQASKKESVAHWEAQIAETIQW 824  
Qy 1031 LTSQKOTMEALKTCTWLEQVNDLALNDELLEKEKQWEAERSVLGDSEKOFECRVERL 1090  
Db 825 VSEKQARGVLOALASQWTE--ELTLRSSLSGSRTLPLMK-VRRSQKLDMSARL-EL 879  
Qy 1091 QRMLDTEKQSRARADORITESTRQVLEAVKHEKAEITLQQAQKLEKLAESLSKNDL 1150  
Db 880 QSALEAE-----IRAKQVHEELRKVKYKDTSLAFESKLKESAK----- 917  
Qy 1151 EKGHMLENNARSLOQKLETERELKRLLEBAQKLOQMDLOKNNHFRUTQGLQZALDRA 1210  
Db 918 -----NRELEEMOSLKRMEEK-----FRADTGL----- 942  
Qy 1211 DLLKTERSDLEYQLENIQVLYSHEKVMGEGTISQOTKLIDFLQAKMDQPAKKKGLFSRR 1270  
Db 943 -----KLPDF----- 947  
Qy 1271 KEDPALPTQVLOYNELKLALEXEKARCALEALEALOKTRIELRSAREEAHRKATDHPHP 1330  
Db 948 -QDPIFE-----YFNAPLAHADHTFTTSSASQOETQASKIDL-----SPSV 987

Qy 1331 STPATARQQIAMSIAI VRSPHEQPSAMSLAPPSSRRKESSTPEFSRRLKERMHNIPHR 1390  
Db 988 SVATSTEQQ---EDAARS-QQRESTVPL---PNTQALAMAGPKPKA-----HQ 1028  
Qy 1391 FNVGLNRRATKCAVCLD-TVHFGRQASKCLECOVMCHPKCSTCLPATCGLPAYATHFTE 1449  
Db 1029 FSIKSPSPQCSHCTSLMVGLIRQYACEVCAFSCHVSCDSAPQVCPIPE----- 1081  
Qy 1450 AFCRDKNNSPGLQKBP-----SSSLHLEGWVKYPRNNKRGQCGQWDRKYIVLEGSV 1502  
Db 1082 -----OSKAPLGVGVORGIGTAYKGVKVPKPT-GVYKQWQAYAVVCDCKLF 1128  
Qy 1503 IYD-NEAREAGQRPVEEFELCPDGDVSIHGAVGASELANATAKAKEA---EADAKLLG-- 1556  
Db 1129 LYDLPEKSTQPGVIASQVLDLRDDEFV-SSVLASDVTHATRDIPCFPRVTASLLGSP 1187  
Qy 1557 ---NSLLKL---EGDDR-----LDNM-----CTLPESDOV--- 1580  
Db 1188 SKTSSLLILTENENKRVKVGILEGLQAILHKNRLRSQVHVHVAQAYDSSLPIKTVLAA 1247  
Qy 1581 -----VLVTEEGLYALNVKNSLTHVPGIGAVFQIYIKDLEKLLMIAGERALCLV 1633  
Db 1248 AIVDGDRIAVGLEBGLVIELTROVIVRAADCKKVQIELAPKEKLILLCGRNH----- 1302  
Qy 1634 DVKKVKSQAQSHLPQPDISPNIPEA-----VKGCHLEGAGKIEGLICCAAMPS 1684  
Db 1303 -----HVHLYPMTSFDCAEASNFIDIKLPETKGCQLIATGLTRKSSSTCLFVAV 1350  
Qy 1685 KWILRYNENLSKYRKETESETPCS-----CIHF-TNYSILITGNKFEIDM 1732  
Db 1351 KRLVLCVEIORTPFPRKNEIYVAPGHVQWAMFKDRLCVGYPGFSLLSIQDGGPDL 1410  
Qy 1733 KQYTLFEBLKNHSLAPAVFAASSNSFPVSIYOVNSAGOREBYLLCFHFVGVFVSYGR 1792  
Db 1411 -----VNPADPSLA---FLSQQSFDAICAVELKS---EYLLCFSHGLYVDPQGR 1455  
Qy 1793 RSRTDCLKWSRLPLAPVREPYLFVTHFNSLEVIEIQARSSAGTPARAYLIDINPRYLGP 1852  
Db 1456 RSRTQELMWAAPVACSCSSHVTIVISEYGVDFVRTME-----WVQTGLARIRP 1507  
Qy 1853 AISSGAIYL 1861  
Db 1508 LNSDGSNL 1516

RESULT 4  
S74244  
serine/threonine-specific protein kinase (EC 2.7.1.-) isoform I, Rho-associated - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 09-Jun-2000  
C:Accession: S74244  
R:Nakagawa, O.; Fujisawa, K.; Ishizaki, T.; Saito, Y.; Nakao, K.; Narumiya, S.  
FEBS Lett. 392, 189-193, 1996  
A:Title: ROCK-I and ROCK-II, two isoforms of Rho-associated coiled-coil forming protein  
A:Reference number: S74244; MUID:96368048; PMID:8772201  
A:Accession: S74244  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1354 <NAK>  
C:Superfamily: unassigned Ser/thr or Tyr-specific protein kinase; protein kinase homolog  
C:Keywords: coiled coil; phosphotransferase; serine/threonine-specific protein kinase  
F:74-338/Domain: protein kinase homolog <KIN>  
F:1229-1283/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 12.08; Score 1261.5; DB 2; Length 1354;

Best Local Similarity 26.18; Pred. No. 4-2e-28;

Matches 397; Conservative 289; Mismatches 537; Indels 297; Gaps 51;

Qy 44 SPLREGILDALFVLFEECSQPALMKIKHVSNFVKYSYDTIAELQELQPSAKDFEVS 103

Db 23 SEVNSDCILLDGLDALVYDLDFPALRKKNKINDFLSRKYKTINKIRDLRMAEDYEVVKVI 82

```
QY 104 GCCHFAEVQVVRKATGDIYAMKVMKKKALLAQEQVSPFEERNILSRSTSPWIPQLQYA 163
Db 83 GRGAFGEVQLVRHKSIRKRYAMKLLSKFEMIKRSDSAFTWEERDIFAPANSPPWVQLFYA 142
QY 164 FQDKNHLXWEEYOPGGDILLSLNRYVEDOLDENLQFYLAELILAVHSHVLMGYVHRDIK 223
Db 143 FQDRRLYVWMEYMPGGDLVNLMSNYD--VPEKWARFYTAEVVLADAJDAHSMGF.IHRDVK 200
QY 224 PENILVDRTHIKLVDFGSAAXNSNMVNAKLPIGTDPYMAPEVILTVNMGSGKTYGID 283
Db 201 PDNMLLDKSGHLKLADFGTCMKVKNKBMVRCDTAVGTDPYISPEVLKSGGD--GYGGR 258
QY 284 CDWWSGVTVAYMIYGRSPFAGTSARTNNIMNQRFKFPDPPKVSDDFLDLTQSLLC 343
Db 259 CDWWSGVFELYELVGDTPFYADSLVGTYSKIMNHKNSLTFFDDNDISKEAKNLI.CAFIL 318
QY 344 GOKERLKFEF---LCHPPFFSKID--WNINRSPFPVPTLKSDDDTNFDPEKSNWVS 398
Db 319 DREVLRGRNGVEIEKHLFFKQDQAWETLRDTAVVPDLSDDIDTNSFDDLEEDKDE 378
QY 399 SS-PCOLSPSGSGBELPVPVGSYKALGILGRSESVSGDSPAKTSMEMKLLIKSKE 457
Db 379 ETFFPI---PKAFVGNQLPVPVGTYSNNRYL-----PSANASENRSNNVDKS 423
QY 458 LODS-QDKCHKMEOEMTRLHRAVSEVAVLQKEVELKASQTQ---SLLBQD-LATYIT 512
Db 424 LQESLQXTTYKLEBQI-----HNEMQLKDEMEQKRTSNLKLKDKIMKELD 468
QY 513 ECGSLKRSLEQARMEVQSDDDKALQLLHDIREQSRKLQEIKEQEYQAOVE-EMRLMMNQL 571
Db 469 EEGNORRNLESA--VSQIEKEMLLQHRINEYQKVEQ--ENEKRNINEVSTLKDQL 523
QY 572 BEDLVASARRSLYSELESRESLAAEFPKPKATECQHKLLKAKDQKGEVEYAKLEKIN 631
Db 524 E-----DURKASQTSOLANEKUTLQKQLEANDLRLTESHTAVLRKSH 568
QY 632 AFQQLKIQLQELKBAKAEERLEKIQNREDSSEGRKIKLVEABERRHSLENKVKRL 691
Db 569 TEMSKSISQL-ESSLNRELQER-NRILENSKSAQDKDYQLQAVLEAERDRG----- 618
QY 692 ETMERENRLKDDIQTK-----SQIQQMADKILEEBEKHRAQVSAQHLVHLKQKQEHY 747
Db 619 -----HDSEMIQDLQARTSLQEEVVKHLKHLNLERVEGERKEAQDMLNHE---KEK---- 666
QY 748 BEKIKVLNDQIKKLADKETTLENMWRHEEAHE---KGKILSQKAMINAMDSKIRSL 803
Db 667 -----NNLEIDLNYK--LKSIOQRLEQVNEHKVTKARLTDKHQSIIEAKSVAMCEM 716
QY 804 EQRIVELSEANKLAANSLFTORNKAQEMIS-ELRQCKFYLETQAGKLEAQRNKLSEEQ 862
Db 717 EKKLAEEREAREKAENRVVETEK-----QCSMLDVLKQSQOKLE-----HLTENKERMED 768
QY 863 LEKISHODSDKNRLELETRLEVSLEHEBQK-----LELKQLTELQLS- 908
Db 769 VKNLALQLEQSNKRLILLQNEUKTQAFADNLKGLKQMKQKEINTLLEAKR-LLEFELAQ 827
QY 909 -----LQERESQTLQARAALESQLRQAKTELEETTAEEABEEIQALTAHRDE 957
Db 828 LTKQYRGNEGQWRELQDLQEAQYESTLYKTQVKELKEEIEEKENRNLKIQELQSEKET 887
QY 958 IQRKPDALNSCTVITDLEEQNLQNTEDNAELNNQNIFYLSQQLDEASGANDEIVOLRSEV 1017
Db 888 LSTQLD-----LAETKAESEQLA--RGILEEQYFELTQESKKAASN----- 927
QY 1018 DHLRREITEREMQLTSQQTWEALTKTCTMLLEQVMDLEALNDELLEKERQWEAARSVLG 1077
Db 928 ---RQEITDKD-----HTVSRLEETNSVLTQKLEMLRKEEELNENRMTAESEYKLK 977
QY 1078 DE-----KSQFECRVRELQRLMDTEKQSRARADORITETSRQVVELAVKHEKAEILALQQA 1132
Db 978 BEEINNLKAAFE-----KNISTERTLKTQAVNKLAEIMNRKDFKIDRKKA----- 1022
```

```
QY 1133 LKEQKLKAESLSDKLNLEKHALEMNARSLOOKLETERE-LKORLLEEQAKLQOQMDL 1191
Db 1023 -----NTQDLRKK-----EXENRKLQLELAQEREKFNQMVVYHQKELN---DM 1062
QY 1192 QKNHIFRLTQGLQALDRADLLKTERSDLEVOLENIQVLYSHEKVMKEGTISQOTKLIDF 1251
Db 1063 Q-----AQLVESC-----THRNELQWLASKESDIEQGRAKLL-DLSDSSTVASF 1106
QY 1252 -----LQAMDQPAKKKKGLFSRRKDEFPALPTQVPLQYNELKLALKEKEKAR 1297
Db 1107 PSADETDGNLPESRIEGLWSVPNRGNIKRYGKKQVYVVSXKLLFYND-----EQDK-- 1159
QY 1298 CAELEBALQKTRIELRSARFEAAHRKATD-----HHPSTPATARQOQIAMSIVRSBEHQ 1352
Db 1160 -----EQSSPMSVLDIDKLFHVRPVTQGDVYR-----AETEEI 1192
QY 1353 PSAMSLAPSSRRKESSTPEEFERRLKERMHNNIPIRENVGNLMRATKCAVCLDTVHFG 1412
Db 1193 PKIQIILYANEGERKDIIEVPVQOGKTNFQNHKGHEFIPITLYHPFANCEACAKPLWHV 1252
QY 1413 ROASKCLIECQVMCHPKC-----STCLPATCGLPABYATHFTFAFCORDKXNSGLQTK 1465
Db 1253 FKPPPALECR-RCHVKCHRDHLKKEDLISPKVSYDVTS-----ARDVLLILACSDEQ 1305
QY 1466 PSSSLHLEGMKVPNNKRG 1485
Db 1306 KKWVTHLV--KKIPKNPPSG 1323

RESULT 5
S70633
N:Alternate names: Rho-associated protein kinase (EC 2.7.1.-), Rho-associated - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 24-Sep-1999
C:Accession: S70633; S77694
R:Matsumi, T.; Amaro, M.; Yamamoto, T.; Chihara, K.; Nakafuku, M.; Ito, M.; Nakano, T.; O
EMBO J. 15, 2208-2216, 1996
A:Title: Rho-associated kinase, a novel serine/threonine kinase, as a putative target for
A:Reference number: S70633; MUID:96208507; PMID:8641286
A:Accession: S70633
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1388 <MATI>
A:Cross-references: EMBL:U36909; NID:gl326077; PID:AA48567.1; PID:gl326078
A:Accession: S77694
A:Molecule type: protein
A:Residues: 1-18;30-34;36-44;58-64;133-140;248-252;291-295;297-305;327-347;350-360;366-37
-1070 <MAT2>
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase C zinc-
C:Keywords: Atg; phosphotransferase; serine/threonine-specific protein kinase
F:90-354/Domain: protein kinase homology <KIN>
F:98-106/Region: protein kinase ATP-binding motif
F:1261-1315/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 11.9%; Score 1253; DB 2; Length 1388;
Best Local Similarity 26.1%; Pred. No. 7.3e-28;
Matches 394; Conservative 291; Mismatches 507; Indels 320; Gaps 50;

QY 15 GAPEPTA---SRASRLNLFQGGPPMTQOQKSPLSREILDALFVLFECSPALMKIK 71
Db 12 GAPEAVSGDGAGASR-----QRKLEALIRDPRIINVESTLLDGLNPLVLDLDFPALRK 66
QY 72 HVSNFVKYKSDTIAELQELQPSAKDPEVRSVLCGCHFAEVQVVRKATGDIYAMKVMKK 131
Db 67 NIDNFLNRYEKIVKIRGLQKMAEDYDVVKVIGRGAFGEVQLVRHKASQKVYAMKLLSKF 126
QY 132 ALIAQSQVFFBEERNILSRSTSPWIPQLQYAFQDKNHLXWEEYOPGGDILLSLNRYED 191
Db 127 EMIKRSDSAFFBEERDIFAPANSPPWVQLFCAFDQDKKLYLMVMEYMPGGDLVNLMSNYD- 185
QY 192 QLDENLIQFLAELIILAVHSHVLMGYVHRDIPENILVDRTHIKLVDFGSAAKNSNM 251
```

186 -VPEKWAFTYAEVVALDAIHSMGLIHEDVKPDNNLLDKHGHLKLADFGTCMKMDTGM 244  
Db  
252 VNAKLPIGTPDYVAPVLTVMNGDGKGTGGLCDWMSGVYIAYEMIGSPAEGETSART 311  
Qy  
245 VHCDDTAGTDPDVISPEVLKSGGD--GYVGRECDWMSGVFLFEMVLGVTPTPYADSLVGT 302  
Db  
312 FNNIMNFORFLKPPDDPKVSSDFLDLIQSLGCKQERLKFEG---LCCHPPFSKIDWN-- 366  
Qy  
303 YSIMPDKHNSLCPEDAEISKIANLICALFDREVLRGNGVEEIKQHPFFKNDQWMD 362  
Db  
367 NTRNSPPVPTLKSDDDTSNDEBPKNW-VSSPCQLSPGSGEELPFGVPSYKAL 425  
Qy  
363 NIRETAAPVVPVPELSSDIDSNFDDIBDDKGVETETPI---PRAFGNQLPFGFTYYREN 419  
Db  
426 GILGRSESVVGLDSPA--KTSMEKKLJKKLEQDSODKCKMEQEMTRLHRRVSEVE 483  
Qy  
420 LLUS-----DSPCKENSIQSR---KNEESQEIQKLYTLEHLS----- 457  
Db  
484 AVLSQKEVELKASQTSRLEQDLATYITECSLSKSLSEARMEVSEQDDKALQLLHDIR 543  
Qy  
458 -----TEIQAKEE---LEQ-----CKSVNTRLEKVAKELEBEITLRKNVESTLR 499  
Db  
544 --FQSRKLOEIKQEOYQAEVEMRMNNOLEBDLYSARERSDLYSSELRESLAAEEFKR 601  
Qy  
500 QLEREKALLQHKNAEYORKADHEADKRRNLENDVNSLKQDLKRRNONSISTE---- 555  
Db  
602 KAPECHQKLLKAKDQKPEVGEYAKLEKINAEQQLKIQ-----ELQEK--LEKAAKER 652  
Qy  
556 KYNQLOQLQDETNALLRTESDTAARLKTQAESSKQIQLESNNRDLQKNCLETAKLK 615  
Db  
653 ARELEKLNREDSGGIRKKLVEABERRHSLNKKVRLTWERENRKLKDIQTSQOI 712  
Qy  
616 LEKEFINLOSVESE---RRDRTHGSEIINDLQGRISGLEEDVYQNGKILLAKVELEKQL 672  
Db  
713 QOMADKILEEKEHREAOVSAQH---LEVHLKQKE-QYEEKIKVLD-NQIKKDL--AD 764  
Qy  
673 Q---ERFTDLEKKNWEDMTYQVKVIOQSLEQETEHAKATKARLADCKNIVESBEAK 729  
Db  
765 KETLENMQSHBEAHEKGI---LSEKAMINAMDSKIRSLERIVELSE-----ANKLA 817  
Qy  
730 SEAMKEVKLSBERTLKQKVENLLAEKRCISLDCDLKQSQKINELLKQDVLNEDV 789  
Db  
818 ANSSL-----FTORMKAQEWISSELROQKPYLETQAGLEAQNRLKEOLEKIS 867  
Qy  
790 RNUTLKIEQTKRCITQNDLRKMTQOVNFKWS-----EKQLKQ-- 829  
Db  
868 HQDHSKRNRLLETLRLREVSLSEHQKLEKRLQTELSQRESQLTALQAARALE 927  
Qy  
830 -----ENNHLKMSL-----EKQNAELRKERQDQGMKQLQDLQEAQYFSTLYK 877  
Db  
928 SQRQAKTELEBTT---AEABEIQALTAHREDEIQKFPALRNSCTVIITDLEQLNOLTE 984  
Qy  
878 TOVRELKEECEETKCKELQKQKQELQDERSLAAQLE-----ITLTAKDSEQLAR 929  
Db  
985 DNAELNNQFYLKQDASGANDVQLRSEVDHLRBITTEREMQLTQSOKTMEALKTT 1044  
Qy  
930 STAEQVSDLEKIMKE-----LEIKEMMARHKQELTEKQATIASLEETNRLTSD 981  
Db  
1045 CTMLEEQVMDLEALNDELLEKQEWANRSLVGLDE-----KSQFECEVRLEQLRMLDTE 1097  
Qy  
982 VANLAN---EKEELNKLKEAQEQ---LSRLKDEEISAAAIKAQPE-----KQLLTE 1027  
Db  
1098 KQSRARADQOITESTRQVVELAVKHEKAEIILALQOALKQKLEASLSDKINDLEKHAML 1157  
Qy  
1028 RTUKTQVANKL-----AEINMRKEPVK-----RGNDTVERK-- 1059  
Db  
1158 EMNARSQQKLETERELKORLLEQAKLQOQMDLQKRNHIFRLTQ---GLQEBALDRADLLK 1214  
Qy  
1060 EKENRKLHMLKSGERE---KLQOMIKYQKELNEMQAIAEESQIRIELQMTLDSKD--- 1113  
Db  
1215 TERSDLEYQLENTQVLY---SHEKVKVQEGTIS---QOQTKLIDF 1251  
Qy  
1114 ---SDIEQLRSQALHIGLSDSSISGGPDTEADGFPESRLGLEWLSLPVRNNTKFGW 1170  
Db

1252 LOAMDQPAKKKKGLFSRRKEDPALPTQV-----PLOYNELKLALEKEKARCAEL 1301  
Qy  
1171 VKXVIVSSKKILFYDSEQKESQNFVWVLDIDKLFHVRPVTQTDVVRADAKEIPRIFQI 1230  
Db  
1302 EEALQKTRIELRSARBEAAHRKATDHPHSTPTATARQQIAMSALVRSPEHQPSAMSLAP 1361  
Qy  
1231 -----LYAN 1234  
Db  
1362 PSSRRKESSTP---BEFSRRLKXRMHNPVGLNMRATKCAVCLDTVHFGQASK 1417  
Qy  
1235 EGSKKEQEFVPEVGEKSNYICHKGHEFTPTLYHF-----PTNCEACMKPLWHMFKPPP 1289  
Db  
1418 CLEQVNCCHKPC 1429  
Qy  
1290 ALECR-RCHIKC 1300  
Db  
RESULT 6  
S69211  
serine/threonine-specific protein kinase (EC 2.7.1.-), Rho-associated - human  
C;Species: Homo sapiens (man)  
C;Date: 06-Dec-1996 #sequence\_revision 06-Dec-1996 #text\_change 24-Sep-1999  
C;Accession: S69211; S71910  
R;Ishizaki, T.; Maekawa, M.; Fujisawa, K.; Okawa, K.; Iwamatsu, A.; Fujita, A.; Watanabe,  
EMBO J. 15, 1885-1893, 1996  
A;Title: The small GTP-binding protein Rho binds to and activates a 160 kDa Ser/Thr prote  
A;Reference number: S69211; MUID:96203110; PMID:8617235  
A;Accession: S69211  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-1354 <ISHJ>  
A;Cross-references: EMBL:U43195; NID:g1276900; PIDN:AA02814.1; PID:g1276901  
A;Accession: S71910  
A;Molecule type: protein  
A;Residues: 187-195;281-288;465-473;573-587;818-828;885-893;934-945 <ISH2>  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase C zinc-  
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
P;74-338/Domain: protein kinase homology <KIN>  
P;182-90/Region: protein kinase ATP-binding motif  
P;1229-1283/Domain: protein kinase C zinc-binding repeat homology <KZN>  
Query Match 11.9%; Score 1250.5; DB 2; Length 1354;  
Best Local Similarity 26.0%; Pred. No. 8.4e-28;  
Matches 393; Conservative 298; Mismatches 543; Indels 275; Gaps 51;  
Qy 44 SPLSREGILLDALFVLFEECSQPALMKIKHYSNFVRKYSDTIAELOEPSAKDPEVRSV 103  
Db 23 SEVNSDCLDGLDALVYDLDFFPALRKNKINIDNLSRYKOTINKIRDMAEDYEVVKVI 82  
Qy 104 GCGHFAEVQVREKATGDIYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPQOYA 163  
Db 83 GRGAFGEVQLVRKSTRKVVAMKLLSKFEMIKSDSAFFWEERDIMAFAFANSPWVQLFYA 142  
Qy 164 FQDKNHYLMEEYOPGGDLSLNRYSDDQDENLIQYLAELILAVSHVLMGVHEDIK 223  
Db 143 FQDDRYLYMMEYMPGGDLVNLMSNYD--VPEKWARFYTAEVVLDAIHSMGFIRHDKV 200  
Qy 224 PENILVDRTHGI KLVDVFGSAAKNNKNNVAKLPIGTPDYMAPEVLTVMNGDGKGTGLD 283  
Db 201 PDNMLDKSHGLKADFGTCMKNNKEGVRCDAVAGTDPDYSPEVLKSGGD--GYIGRE 258  
Qy 284 CDWMSGVYAVEMTYGRSPAEGTSARTFNINMFORFLKFPDPPKVSSFDLIQSLLC 343  
Db 259 CDWMSGVYAVEMTYGRSPAEGTSARTFNINMFORFLKFPDPPKVSSFDLIQSLLC 318  
Qy 344 QKRELKFEK---LCCHPFFSKID--WNINRNSPPFPVPTLKSDDDTSNDEPEKMSWVS 398  
Db 319 DREVLGRNGVEETKRLHFFKNQDQWAMETLDTVAVVVPDLSDDIDTSNDDLEEDKGE 378  
Qy 399 SS-FCQSPSGFSGEELPFVGFYSYKALGILGRSESVVGLDSPAFTSSMEKLLIKSKE 457  
Db 379 ETTFPI---PKAFVGNQLPFGVFTY-----YSNRRYLLSSANPDNRTSSN-----ADKS 423

QY 458 LQDS-QDKCHKMEQETRLHRRVSEVAVLSQKEVELKASQETOR---SILEOD-LATYIT 512  
Db 424 LQESLQXTYKLEQI-----HNEMQLKQMEQKCHTSNKLKDKWKELD 468  
QY 513 ECSSLKRSLEQARMEVQSDDKALQLLHDIRBQSRKLQEIKEQEYQAQVE-EVRLMMNL 571  
Db 469 EECNQRNLEST---VSQIEKEMQLQHRINEYQRAEQ--ENEKRNVENEVSTLKQOL 523  
QY 572 EEDLVASARRSDLYESLRESLAAREFKRKATECQHKLLKAKQDQKPEVGEYAKLEKIN 631  
Db 524 -EDLKVSQNSQL-----ANE---KUSQLOQEEANDLLRTESDFAVLRKSH 568  
QY 632 AEQQLKIQLOEKLEAKAERARELEKLQNRSDSEGIRKLVAEAEERHSLNKKVRL 691  
Db 569 TEMSKSISQL-ESSLNRELQER-NRILENSKSTQDKDYQLOALAEARERDRG----- 618  
QY 692 ETWERENLKODIQTK-----SQOIQOMADKILEEKKHREARQVSAQHLEVLKQKEQHY 747  
Db 619 -----HDSMIGLOARITSLQEBEVKHLKHLKVEGERKEAQDMLNHE---KEK--- 666  
QY 748 EEKIKVLDNQIKKDLADKETLENMQRHEEEAHE-----KGKILSEQKAMINAMDSKIRSL 803  
Db 667 -----NNLEIDLNYK--LKSLOQLQEOVNEHKVTKARLTDKHQSTEEAKSVAMCEM 716  
QY 804 EGRIVELSEANKLAANSLSFTORNKKAQEMIS-ELRQOKFYLETQAGKLEAQNRKLEQ 862  
Db 717 EKKLEEREAREKAENRVV---QIEKQCSMLDVLKQSQQKLEHLTGKE-----RMEDE 768  
QY 863 LEKISHODSDKNRILLELETRLEVSLEHEEOK-----LELKROLTELQLS- 908  
Db 769 VKNLTLOEQENKRLLLQNEKLTQAFADNKLKLEKQKQBNINTLLBAKR-LLSEFEAQ 827  
QY 909 -----QERESQJLTALQAAPALESQRQAkteLEETTAAEIEEIQALTAHRDE 957  
Db 828 LTKQYRGNEGQRELQDQAEQYFSTLYKTQVKELKEIEBKRNENLKIKIQELQNEKET 887  
QY 958 IQRKEDALENSTVITDLEEQNLQNTEDNAELNNQNFYLSKQLDASGANDBIQVLRSEV 1017  
Db 888 LATQID-----LAETKAESQJLA--RGLLEEQYFELTQSSKKAERN----- 927  
QY 1018 DHLRREITERMQLTQSQTMEALTKTCMTLBEQVMDLEALNDELLEKXERQWEAQRVILG 1077  
Db 928 ---RQEIITDKD-----HTVSRLEANSMLTKDIEILLRRENEEITKXKKAEEYKLEK 977  
QY 1078 DE-----KSQFCRVRELQMLDTPKQSPARADQRTSQRQVVELAVKHEKAEIILAQQA 1132  
Db 978 EEEISNLKAAFP-----XNINTERLTQAVNKLAEIMNRKDFKIDRKKANTQDLRKK 1030  
QY 1133 LKEQK---LKAESLSDKLDLLEKXHAMLEMNARSLOQKLETERELKORILLEQA---KLO 1186  
Db 1031 EKENKLOLELQNEQREKFNQVVKH-----QKEL---NDMQQLVECAHNRNELQ 1077  
QY 1187 QQMDLQKNHIFLPTQGLQALDRADLLKTERSDLEYQLENTQVLYSHEKVKXEGTISQOT 1246  
Db 1078 MOLASKESDIEQLRAKLKLDSDSTSVASFPSSAD-----ETDGNLPSER 1120  
QY 1247 KLIDFLQAKMDOPAKKKGLFSRRKEDPALPQVPLQVYNELKLALEKAKCAELEELAQ 1306  
Db 1121 -----IEGWSLVNPGNTIKRYGKKQYVWVSKLIFYNDEQCKEQSNPWSVLOIDKIL- 1173  
QY 1307 KTRIELRSAREBAARHKATDHPHPSTPATARQOIAMSAIVRSPEHQPSAMSLAPPSSRR 1366  
Db 1174 ---FHVPRVPTQGVYRAET-----EIEPKIFQILYANEGEC 1206  
QY 1367 KESSTPEFSRRLKERMHNIHPRNVGNLNPATKCAVCLDTHVHGRQASKLCQCMVCH 1426  
Db 1207 RDXVEMEPVQQAETNFQNHKGHEFIPTLYHFPANCDAKAPLMHVFVPPPALECR-RCH 1265  
QY 1427 PKCS-----TCLPATCGLPAEYATHFTFAFCRDKMNSPGLQTKPEPSSSLHLEGW 1476  
Db 1266 VKCHRDHLDKKEDLIC-----PKVSYDYVTSA--RDMILLASQDEQKKWTHLV--K 1314

QY 1477 KVERNNKRG 1485  
Db 1315 KIPKNPPSG 1323  
RESULT 7  
S74245  
serine/threonine-specific protein kinase (EC 2.7.1.1-) isoform II, Rho-associated - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 24-Sep-1999  
C:Accession: S74245  
R:Nakagawa, O.; Fujisawa, K.; Ishizaki, T.; Saito, Y.; Nakao, K.; Narumiya, S.  
FEBS Lett. 392, 189-193, 1996  
A:Title: ROCK-I and ROCK-II, two isoforms of Rho-associated coiled-coil forming protein s  
A:Reference number: S74244; MUID:96368048; PMID:8772201  
A:Accession: S74245  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1388 <NAK>  
A:Cross-references: EMBL:U58513; NID:g1514697; PIDN:AAC53133.1; PID:g1514698  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase C zinc-  
C:Keywords: phosphotransferase; serine/threonine-specific protein kinase  
F:90-354/Domain: protein kinase homology <KIN>  
F:1261-1315/Domain: protein kinase C zinc-binding repeat homology <KZN>  
Query Match 11.9%; Score 1245; DB 2; Length 1388;  
Best Local Similarity 25.8%; Pred. No. 1.2e-27;  
Matches 381; Conservative 296; Mismatches 517; Indels 284; Gaps 45;  
QY 32 QGKPPFTQQQMSPLSREGILDALFVLFERCSQPALMKIKHVSNFVKYSDTTAELQELQ 91  
Db 27 QRLKALIRIPRSPINVESLDDGLNSLVLDLPFAKKNKINIDFLNRYKIVKIRGLQ 86  
QY 92 PSAXDFEVRSLVGGCHFAEVQVVRKATGIYAKVMKXKALLAQEQVSPFEERNILSR 151  
Db 87 MKAEDYDVXVVGAGGEVQLVVRHKASQVYAMKLLSKEMIKRSDSAFFWEERDIMP 146  
QY 152 STSWIQLQVAFQDKXNHLIMEEYQGGDLLSLLNRYEDQDLNLTQFYLAELILAVHS 211  
Db 147 ANSWVWVQLFCAFODDRYLYVMVMEYMPGGDLVNLMSNYD--VPEKWAQFYTAEVVLA 204  
QY 212 VHLNGYVHRDIPKPNILVDRTHGHLKLVDFGSAKMSNKVNAKLPITGTDYMAPEVLTV 271  
Db 205 IHSGLIHRDVKPNMLLDXGHLKLAFTGTCMKNDTGMVHCDTAVGTTPDYISPEVLKS 264  
QY 272 MNGDGKTYGLDCDWSVGVIAEYMIYGRSPFAEGTSARTFNIMNPFQFLKPPDPKVS 331  
Db 265 QGGD--GYGREGCDWMSGVVLFEMLVGDTPFVADSLVGTYSKIMDKHNSLCFPDETEIS 322  
QY 332 SDFLDLTQSLCCQKELKTEG---LCCHPFPFSKIDWN--NIRNSPPFPVPTLKSDDDTS 386  
Db 323 KHAKNLICAFUTDREVLGRNGVEETKQHPFFKXNDQMNWDNIRETAAFPVPELSSD-DSS 382  
QY 387 NFDBPEKNSW-VSSSPQQLSPGSGFGBELPFVGFYSKALGILGRSESVVSGLDSP--AK 443  
Db 383 NFDDIEDDKGDETFFPI---PKAFVGNQLPFIQFTYFRENLLLS-----DSPCRE 430  
QY 444 TSMNEKLLTKSELQDSQDKCHMQEMTRLHRRVSEVAVLSQKEVELKASQETORSL 503  
Db 431 NDAIQTR---KSESOBIQKKYALEEHLs-----SEVQA---KEELEOK----- 469  
QY 504 EQDLATVITECSLSKRSLEQARMEVSGEDDKALQLLHDIR--EQRKLOEIKQEYQAQV 561  
Db 470 -----CKSINTFLEKTAKLEBEETILRSKSVSTLRQLEREKALLOKHNAEQVKA 519  
QY 562 EEMRLMNQLLEEDLVASRRSDLYESLRESLAAREFKRKATECQHKLLKAKQDQKPEV 621  
Db 520 DHEADKRNLENDVNSLKQLEDLKKRNQSSQISTE---KVNQLQQLDEANALLRTES 575  
QY 622 GEYAKLEKINAEQOLKIQ-----ELQEK--LEKAAKERAERELEKIQNRSDSSGIRK 672  
Db 576 DTAARLURTKQAESSKQIQQLLESNNRDLQDKNCULETAKUKLEKXEFINLOSALES----- 629



```
Db 671 IKKLK-----QRMDERASHTAQSEQEMKQLEAHYERAQKM-----L 707
QY 976 EEQLNQLTBDNAELNNQNFYLSKQLD--BAGANDEIVQLRSEVDFHLRRE-ITEREMQIT 1032
Db 708 QDNVEQMVNENRGLRDEIEKLQSQMAALPRGGLNEQ--QLHFIENWVSEERATREMENL 765
QY 1033 SQKQT--MEALKTCTMLEBEQWDLAEALNDELLEKEKRWQWARSVLGDEKSFQRCVRBEL 1090
Db 766 TKKITGEVSESLKNNGLTTSNYI-----QNTPSGWS-----RRMNV 803
QY 1091 QRM--LDTEKQSRARADORITESQVVELAVXEHKAEILALQALKEOKLKAESLSDKLN 1148
Db 804 ARKDGJLQROQLQABIDAK-----LKLKAEELKNSGEQVLTSAARDL 844
QY 1149 DEKKHAMLENNARSLOQKLETERELKORLLEEQAKLQOQMDLQ--KNHIFELTQGLQGBA 1206
Db 845 DTEKRMA-----SLMEVAMLKQOKNIENSDDSAFSTWG----- 879
QY 1207 LBRADLLKTERSDLEYOLENIQVLYSHEKVXVEGTIS-----OOTKLIDFLQAKM 1256
Db 880 --RGDLMISWND--YEMSN--SSLRQEMISQSTPSYENAILLHDHQVPRKVDLRYK- 933
QY 1257 DOPAKKKGLFRRKXEDPALPTQVPLQYNELKLALEKAKARCAELEALQKTRIELRSAR 1316
Db 934 QXPMKTASGIFS----- 945
QY 1317 BEAHRKATDHPSTPATARQOIAWSAIVRSPHQPSNMSLLAPPSSRRKESSTPEBFS 1376
Db 946 -----PVSTISAMERG--HNFERMKIKTP----- 966
QY 1377 RBLKERMHHNI PHRFNVGLNMRATKCAVCLD--TVHFGROASKCLCQVMCHPKSTCLPA 1435
Db 967 -----TKCGHCTSLIGLDROGLFCQSCQVACHVSCAERSQ 1003
QY 1436 TCGLPAYATHTEAPCRDKMNSPGLQTKPESSSLHEGWMKVPNNRKGQGWDRKXIV 1495
Db 1004 SCPVPEE----- 1020
QY 1496 LEGSKVLIYDNAREAGQRPVEEFLC-LPDGDVSIHGAVG-----ASE 1538
Db 1021 AEQNRKQDVKNRILRLVDMRDPDFVCGVSEADV-IHAQKGDIPIKIFRVTITQILNSSSE 1079
QY 1539 LANTAX-----AEKAEADAK-----LLGNS-----LLKLEGD-----DRL 1568
Db 1080 YSSSGKFYTLFAETEERKKRWVALSELKTLRRSKLADRKAFLVKEVDFVTLTPSIRV 1139
QY 1569 DMNCTLPFSDQVLYGTEGLVNLVNLKSLTHVPGIGAVFQIYIILKLEKLLMAGGER 1628
Db 1140 AQCCAIIDRSKIVGFSDHGLYCIETISROLLIPVG-----EK-----ENKQ 1182
QY 1629 ALCLVDVKKVQSLAQSHLPAQP-----DISPNIFFAVKGCFLFGAGKIE 1674
Db 1183 CVETVEYDEAEQLLMWIVGPAKDRHVRIVPSAALDGRDLKWKVNDTKGCHLLAVGTNNP 1242
QY 1675 G---LCICAMPKSVILRYNENLSKYCIKRIETSEPCSCITHFNYSILIGTNKFYEI- 1730
Db 1243 GGRAGFFAVAFKSVTIFQIDRSKRHKWKDLAMPGTPOSTAIENGRLYVGFSSFRS 1302
QY 1731 -----DMKQYTL--EFLDKNDHSLAPAVFAAASSNPVSIVOVNSAGOREE 1775
Db 1303 SLVGVDSPPVSGDASGAVLQHSILVNMDTSLQ--FLNQOTSTEAKLI-VNVPSPDE 1358
QY 1776 YLLCFHFQVFDVSGRRSRDLDLKWRLPLAFAYREPLYFVTHFNLSLEVII 1828
Db 1359 YLLVFNMIGLYVNMGRSRLPEVFPPTQAKYFAYHEPVLVFSNEVDIFNV 1411
```

## RESULT 9

T25539

hypotheoretical protein C10H11.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-Jan-2000

C:Accession: T25539

R,Dante, M.; Wamsley, P.  
submitted to the EMBL Data Library, February 1997  
A:Description: The sequence of *C. elegans* cosmid C10H11.  
A:Reference number: Z20047  
A:Accession: T25539  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1173 <DAN>  
A:Cross-references: EMBL:U08311; PIDN:AAB42348.1; GSPDB:GN00019; CESP:C10H11.9  
A:Experimental source: strain Bristol N2; clone C10H11  
C:Genetics:  
A:Gene: CESP:C10H11.9  
A:Map position: 1  
A:Introns: 23/3; 51/3; 104/3; 343/1; 478/2; 868/3; 891/3; 970/3; 1027/3; 1114/1  
C:Superfamily: hypotheoretical protein C10H11.9; protein kinase homology

```
Query Match 10.6%; Score 1112.5; DB 2; Length 1173;
Best Local Similarity 29.9%; Pred. No. 4.6e-24;
Matches 291; Conservative 210; Mismatches 346; Indels 127; Gaps 33;

QY 44 SPLSRGILDALFVLFECSQPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVSILV 103
Db 15 SPINIESLDTITALTVDCKIPVLMRMKSVDNFISRYERVVESLAALRMKAADFQKLVI 74
QY 104 GCGHFAEVQVREKATGDIYAMKVMKKALLAQBOVSFFEBEERNILSRSTSPWIPQOYA 163
Db 75 GRGAFGEVHLVRHTRTNTVYAMKWLKDDMIKRADSAPFWEERDIMAHANSEWILRLOYA 134
QY 164 FQDKNHLVMEYOPGGDLLSLNRYEDQDENLIQYLAELILAVHSVHLMGYVHRDIK 223
Db 135 FQDPRHLYMWMEYMPGGDLVNLMTSYE--VSEKWTREYTAIVEALAAALHSMGYIHRDVK 192
QY 224 PENILVDTGHIKLVDFGSAAKNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKGTGLD 283
Db 193 PDNMLIISIGHIKUADFGTCVKNANGVVRCSTAVGTPDYISPEVLNRNQODAB--FGKE 250
QY 284 CDWMSVGVIAYEMTYGRSPFAEGTSARTFNINMNFQFLKFPDDPKVSSDFLDLIQSLLC 343
Db 251 VDWMSVGVIYEMLVGTPEVAEALVSTYTNIMNKHSLAKFPDPLISTQAKDIKKFLS 310
QY 344 GQKERL--KFEGLCCHPFFSKIDWN--NTRNSPPPPVPTLKSDDDTSNTDEPEKNWS 398
Db 311 AAPRLGRNSVDDIRNKHFFVNDWTEPATLRASPPVPSLKSDDDTTFEEIETRDNR 370
QY 399 SSPQLSPSFGSGBELPFVGFYSKALGILGRSESVSGLDSPAKTSSME--KKLLKSK 456
Db 371 AGDFQL-PKTFNGQLPFIQFTSYNEYSV--KNLLKHGAGSKQNGIEQHKPQTVEQ 426
QY 457 ELQDS-----QDKCHRMQEMLRHRVSEVAVLSQKEVELKASQTORSLLLEQDLATY 510
Db 427 PLTNHASGVPEEKYEAVKMLDKNREFELLKDSIARNEIRAKMIENKNSLS----- 480
QY 511 ITECSSLKRSLEQARMEV--SQEDDKALQLLHDIREQSKLOEIKEQEYQAOVBEELM 567
Db 481 -TKISDLERELKONKDLRHGADSDAKVNELAVELR-----MSKEYNSESELSKPRDK 534
QY 568 MNQLEEDLVARSARDLYESELRESRL-----AAEFKR-----KATECOHKL 610
Db 535 CEQLKEDL--RKXGSLAQEKNETQVFOQKXDADEAFAEIKREDYELLQTRENEKSVQL 591
QY 611 LKADQCKPEVGEYAK--LEKINAEQLKIQELQEKLEKAAKAPARELEKIQNRESSE 668
Db 592 KKALDERK-ENGAYQQSVAKATDAEWERKQVFEYKQLEHANDER-KREEQRTAAEPDQS 649
QY 669 GIRKKLVEABERRHSLENKVRL-ETMERRENKDKDIQTKSQ--QIQOMAD-----K 718
Db 650 RVARKLAGIEANYELLQNDYKSMKEARKDLERDQDVITEKRRIEIRVEQLMDSRNTDER 709
QY 719 LLEL-BEKHREAQVSAQHLVHLKQK---BOHYEEKIKVLNDQIKKDL--ADKE----- 766
Db 710 VLSLQDELVESQFEAKYKEDGLRGIDGFKHELENEKMKQTLEENLLVADKXERGLMK 769
QY 767 TLENMQRHEEAHEKGIILSEQAMINAMDSKIRSLEQRIVELSEANKLAANSLSLFTQ 826
```

```

Db      770 EVQELMQRHWEITNK-----DQTLKHELTQDLEIKQSKIESSEQ-----ES 812
QY      827 NMKAQEMISELQQOQFYLEYTOAGKLEAQNKRLEEQLEKISHQDSD-----KNRLLELET 882
Db      813 NDK---QTIADLRKK---LDLEKHAHKAVINKLEEMAKRQPLKGEKGVTSALIKKER 866
QY      883 RLREVSLEHEQKL-----ELKRLQTELOLSQERESQITLQAARAALESQRLQA 933
Db      867 EI--MALEQERDMSKRIALFAYENDKQAEHFNIAQD-----MOTTQDALRELKEC 917
QY      934 KTBLETTAEABEE 947
Db      918 KEHLARNVNTRYE 931

RESULT 10
B49364
protein kinase (EC 2.7.1.37), myotonic dystrophy-associated - human
N/Alternate names: myotonin protein kinase
C/Species: Homo sapiens (man)
C/Date: 20-Feb-1995 #sequence, revision 20-Feb-1995 #text change 24-Sep-1999
C/Accession: B49364; S48681; A42101
R/Shaw, D.J.; McCurrach, M.; Rundle, S.A.; Harley, H.G.; Crow, S.R.; Sohn, R.; Thirion,
Genomics 18, 673-679, 1993
A/Title: Genomic organization and transcriptional units at the myotonic dystrophy locus.
A/Reference number: A49364; MUID:94140369; PMID:7905855
A/Accession: B49364
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-624 <SHA>
A/Cross-references: GB:U19268; NID:G307176; PIDN:AAA36206.1; PID:G307177
R/Sasagawa, N.; Sorimachi, H.; Maruyama, K.; Arahata, K.; Ishiura, S.; Suzuki, K.
FEBS Lett. 351, 22-26, 1994
A/Title: Expression of a novel human myotonin protein kinase (MTPK) cDNA clone which enc
A/Reference number: S48681; MUID:94357271; PMID:8076686
A/Accession: S48681
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-417, 'L', 419-468, 'P', 470-544, 'MAPRPWLASARWNGQAPCTAATCCSLPGS', 'LGAYRRRPFC
A/Cross-references: GB:S72883; NID:G633864; PIDN:AAB31800.1; PID:G633865
R/Brock, J.D.; McCurrach, M.E.; Harley, H.G.; Buckler, A.J.; Church, D.; Aburatani, H.;
S.; Davies, J.; Shelbourne, P.; Buxton, J.; Jones, C.; Juvonen, V.; Johnson, K.; Harger,
Cell 68, 799-808, 1992
A/Title: Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repea
A/Reference number: A42101; MUID:92154692; PMID:1310900
A/Accession: A42101
A/Molecule type: mRNA
A/Residues: 'PPGQVGRGLLAVGGA', 57-550, 'WLWASARWNGQA', 564-624 <BRO>
A/Cross-references: GB:M94203; GB:M91465; NID:G186755; PIDN:AAA64884.1; PID:G186756
A/Note: sequence extracted from NCBI backbone (NCBIN:82640, NCBI:82641)
C/Genetics:
A/Gene: GDB:DM
A/Cross-references: GDB:119097; OMIM:160900
A/Map position: 19q13.2-19q13.3
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; phosphotransferase
F:69-339/Domain: protein kinase homology <KIN>
F:77-85/Region: protein kinase ATP-binding motif

Query Match      8.4%; Score 878.5; DB 2; Length 624;
Best Local Similarity 37.5%; Pred. No. 6.9e-18;
Matches 194; Conservative 104; Mismatches 178; Indels 41; Gaps 11;

QY      46 LSRREGILDALFVFEBCSQPALMKIKHVSFVRKYSDTIAEQELQPSAKDFEVSRLVGC 105
Db      20 LGEPLDLLLVGHQELGASLQADKYVADFQWAEPIVVRLEKVELQDRDDEILKVIQR 79
QY      106 GHFAVOVVRKATGDIYAMKMKKALLAQEQVSPFEERNILSRSTSPTPQLOVAFQ 165
Db      80 GAFSEVAVVRKQTGVYAMKMKKALLAQEQVSPFEERNILSRSTSPTPQLOVAFQ 165
QY      166 DKHLYLMEEYQPGGDLTLLNRYEDQDLENLIQFYLAELILAVSHVLMGVVHRDIKPE 225

```

```

Db      140 DENLYLWMEYVGGDLTLLSKFGERIPAEAMARFYLAETVMAIDSVHRLGYVHRDIKPD 199
QY      226 NILVDRTGHIKLVDGSAAKMNSKMYNAKLPGTDPYMAPEVL-TVMNGDGKGTGLDC 284
Db      200 NILDRGCHRLADDFGCLKLADGTVRSILVAVGTPDYLSEILOAVGGGPGTSGVPEC 259
QY      285 DWMSGVGIAYEMIIYGRSPFAEGTSARTFNIMNFQRLKFP-DPPKVSSDFDLIOLSLC 343
Db      260 DWALGVFAVEMFYGTPTFVADSTAEYGIKVHKHLSPLVDEGVPEEARDFIQRLLC 319
QY      344 GQERLAFEG---LCCHPFFSKIDMNNIRNSPPFVPTLKSDDDTNFDPEPKNSWSSS 400
Db      320 PPETRLRGGGAGDFRTHPFFFGDWDGLRDSVPFFPDFFEGATDTCNFDLVEDGLTAMET 379
QY      401 PCQLSPSGFSGEELPFVFGYSKALGILGRSESVVSGLDSPAKTSMKKLL---INKE 457
Db      380 LSDIREGAPLGVLHPFVGYSYS---CMALRDSEVPG---PTPMEVEAEQLLEPHVQAPS 432
QY      458 LQDS---QDKCHKM-----EQEMTLHRRYSEVAEVLQSEV--ELKASETQRS 501
Db      433 LEPSVSPQDETAEVAVPAAPAAEAEBVTLRLEQALEEVLTRQSLRSEMAIRTD-- 490
QY      502 LLEQLATYITECSSLKRSL-----QARMEVVSQED 532
Db      491 --NQNFASQLREAFARNEDLEAHVRQLQERMEQLQAE 525

RESULT 11
S71829
serine/threonine-specific protein kinase (EC 2.7.1.-) - mouse
N/Alternate names: myotonic dystrophy-associated protein kinase; myotonin protein kinase
C/Species: Mus musculus (house mouse)
C/Date: 06-Dec-1996 #sequence, revision 06-Dec-1996 #text change 28-May-1999
C/Accession: S71829; S38815
R/Mahadevan, M.S.; Anemiyi, C.; Jansen, G.; Sabourin, L.; Baird, S.; Neville, C.E.; Worms
submitted to the EMBL Data Library, January 1993
A/Description: Structure and genomic sequence of the myotonic dystrophy kinase (DM kinase
A/Reference number: S71829
A/Accession: S71829
A/Molecule type: DNA
A/Residues: 1-557 <MAH>
A/Cross-references: EMBL:Z21505
R/Mahadevan, M.S.; Anemiyi, C.; Jansen, G.; Sabourin, L.; Baird, S.; Neville, C.E.; Worms
Hum. Mol. Genet. 2, 299-304, 1993
A/Title: Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene.
A/Reference number: S38815; MUID:93271990; PMID:8499920
A/Accession: S38815
A/Molecule type: DNA
A/Residues: 1-53 <MAW>
A/Cross-references: EMBL:Z21503
C/Genetics:
A/Introns: 54/1; 84/3; 112/3; 144/3; 194/2; 225/3; 294/3; 382/3; 411/2; 448/3; 503/2
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; phosphotransferase; protein kinase
F:69-339/Domain: protein kinase homology <KIN>
F:77-85/Region: protein kinase ATP-binding motif

Query Match      8.2%; Score 862; DB 2; Length 557;
Best Local Similarity 36.6%; Pred. No. 1.8e-17;
Matches 199; Conservative 104; Mismatches 191; Indels 50; Gaps 13;

QY      46 LSRREGILDALFVFEBCSQPALMKIKHVSFVRKYSDTIAEQELQPSAKDFEVSRLVGC 105
Db      20 LGEPLDLLLVGHQELGASLQADKYVADFQWAEPIVVRLEKVELQDRDDEILKVIQR 79
QY      106 GHFAVOVVRKATGDIYAMKMKKALLAQEQVSPFEERNILSRSTSPTPQLOVAFQ 165
Db      80 GAFSEVAVVRKQTGVYAMKMKKALLAQEQVSPFEERNILSRSTSPTPQLOVAFQ 165
QY      166 DKHLYLMEEYQPGGDLTLLNRYEDQDLENLIQFYLAELILAVSHVLMGVVHRDIKPE 225
Db      140 DENLYLWMEYVGGDLTLLSKFGERIPAEAMARFYLAETVMAIDSVHRLGYVHRDIKPD 199

```



```
Db 1127 GESTKSIYQH---STLLELDQDKQAKYFLIPAMLNBPAA-----SRLMKRGKGLHIY 1177
Qy 1387 IPRFNVGLNRR-ATKCAVCLDTVH--FGRQASKLEQVMCHPKSTCLPATCGLPDABY 1443
Db 1178 NDHTF-VAVXVKGATCNVQQIRSFSSQAYQCRDCRWCHKTC-----1222
Qy 1444 ATHF-TEAPC 1452
Db 1223 --HYKTDAPC 1230

RESULT 13
G86431
protein kinase Ts18.9 protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 19-Jul-2002
C/Accession: G86431
R/Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: G86431
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-522 <STO>
A/Cross-references: GB:AE005172; NID:G4587520; PIDN:AAD25751.1; GSPDB:GN00141
C/Genetics:
A/Map position: 1
C/Superfamily: Arabidopsis thaliana protein kinase Ts18.9; protein kinase homology

Query Match 6.8%; Score 712; DB 2; Length 522;
Best Local Similarity 35.7%; Pred. No. 2.2e-13;
Matches 157; Conservative 82; Mismatches 133; Indels 68; Gaps 9;

Qy 69 KIKHVSFVRKYSDTTAELOQLPSAKDFEVRSLVCGHFAEVQVVRKATGDIIVAKVM 128
Db 93 KMDILKNFEKMEYN-RIQRQXGVDDFELLTMIGKAFGEVRVCRKSTGSVAMKRL 151
Qy 129 KKALLAQEQVSFEERNILSRSTSPWIPOLQYAFQDKXNHLVLMEEYOPGGDLSLLNR 188
Db 152 KKSEMLRRGQVHVKAERNVLAEDVSPFIVKLCYSQDDDEHLVLMELYPGGDMTLLMR 211
Qy 189 YEPOLDENLIOFLAELLIVHSLMGVYVHRDIKPENLIVDRTHIKLVDFGSAKWN 248
Db 212 -KDTLREDETRFYVAQTILAIETSIKHNYIHRDIKPDNLLITRNHGIKLSDFGLSKLES 270
Qy 249 -----SKMNAKLPIGTPDYMAPEVLTVMNGDG 276
Db 271 KNFPDFKAEIVDRSTKPAAEHRLSKPPSPRRTQEQQLLHWQNRRTLAFSTVGTPTYI 330
Qy 265 APEVLTVMGDGGKGTGLDCDWSVGVIAEYMIYGRSPFAECTSARTFNNIMNFORFLKP 324
Db 331 APEVL-----KKGVMGECMDWSLGAIMEFVLGVPPFPPEPLATCRKIYNWTKLKF 384
Qy 325 PDPKVSDFDLIOQLCGQERLKFEG---LCCHPPFSKIDWNIRNSPPFPVPTLKS 381
Db 385 PDEAKLSIEVKLIRLLCNVQRLGTGKGVHKAHPFRGVENERLYESNAPYTPQVXH 444
Qy 382 DDTSNFDPEPKNSVSSPCQLSPSG-----FGSEELPFVGFYSKALGILGRSESV 435
Db 445 ELDTQNFEPDE-----VPSTCTSSKSPWRKMSKDNANFLGYTF-KNLEIV--DEHI 497
Qy 436 SGLDSPAKTSMEKLLIKS 455
Db 498 PGVAELKRKSKTANKPSLKT 517
```

## RESULT 14

```
S42864
protein kinase (EC 2.7.1.-) - common ice plant (fragment)
C/Species: Mesembryanthemum crystallinum (common ice plant)
C/Date: 06-Jan-1995 #sequence_revision 23-Jul-1999 #text_change 19-Jul-2002
C/Accession: S42864
R/Baur, B.; Winter, K.; Fischer, K.; Dietz, K.
submitted to the EMBL Data Library, March 1994
A/Description: Molecular cloning and characterization of several protein kinases from pla
A/Reference number: S42864
A/Accession: S42864
A/Molecule type: mRNA
A/Residues: 1-479 <BAU>
A/Cross-references: EMBL:Z30329; NID:G457688; PIDN:CAA82990.1; PID:G457689
C/Superfamily: Arabidopsis thaliana protein kinase Ts18.9; protein kinase homology
C/Keywords: ATP; phosphotransferase; protein kinase
F:38-338/Domain: protein kinase homology <KIN>
F:46-54/Region: protein kinase ATP-binding motif

Query Match 6.7%; Score 706.5; DB 2; Length 479;
Best Local Similarity 35.18; Pred. No. 2.9e-13;
Matches 163; Conservative 81; Mismatches 164; Indels 57; Gaps 8;

Qy 76 FVRKYSPTIAELOQLPSAKDFEVRSLVCGHFAEVQVVRKATGDIIVAKVMKKALLA 135
Db 19 FLEKETEYMLRQHKMGADDFELLTMIGKAFGEVRVCRKKTGTHVYAMKLLKXSEMLR 78
Qy 136 QEOVSFTEEBERNILSRSTSPWIPOLQYAFQDKXNHLVLMEEYOPGGDLSLLNRYEDQDLE 195
Db 79 RGQVHVKAERNLAEVDVSNICVLYCSPQDEEYHLYMEYLPGGDMTLLMR-KOTLIE 137
Qy 196 NLIQFYLAELILAVHSLMGVYVHRDIKPENLIVDRTHIKLVDFGSAKWN-----247
Db 138 DEARFYVAETILAIETSIKHNYIHRDIKPDNLLDKFGLRLSDFLGKLPDCLDSTLEEKD 197
Qy 248 -----SKMNAKLPIGTPDYMAPEVLTVMNGDG 276
Db 198 FEVNGNGGSPNENSGTKPRRTQEQQLQHWQNRMLAYSTVGTPTYIAPEVLL-----251
Qy 277 KGTGLDCDWSVGVIAEYMIYGRSPFAECTSARTFNNIMNFORFLKFPDPPKVSDFLD 336
Db 252 KKGVMGECMDWSLGAIMEYMLGVPPFPYSDPDMSTCRKIYNWTHLKFPEEAKLSPEAKD 311
Qy 337 LIQSLCGQERLKFEG---LCCHPPFSKIDWNIRNSPPFPVPTLKSDDDTNFDPEPK 393
Db 312 LISKLLCNVTQRLGSGNKAHEIKLHPWFNGIDWERIYQMEAAFIPEVNDDELDTQNFKEFEE 371
Qy 394 NSWSSSPCOLSP--SGFSGEELPFVGFYSKALGILGRSESVVSGLDSPAKTSMEKKL 451
Db 372 ADNSQSSTKAGPWRKMLSKDLNLFVGYTY-KNFEIV--NDYQVPGIAELKKDTPKRP 428
Qy 452 LIKS---KEIQDSQDKCHKMEQEMTLHRRVSEVAVLQSEVEL 493
Db 429 SIKSLFEDESSDSEATSGDQVQSGFNLNLLPQLLEVSTQTEV 473
```

## RESULT 15

```
T26101
hypothetical protein W02B8.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T26101
R/Sims, M.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z20151
A/Accession: T26101
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1256 <WIL>
A/Cross-references: EMBL:Z81136; PIDN:CAB03458.1; GSPDB:GN00020; CBSP:W02B8.2
A/Experimental source: clone W02B8
```

Qy	1463	TKPSSSLHLEGHMKYPRNNKGQOQWDRKYI1JLGSKVLIYDNEAREAGORPVEE----	1518
Db	812	-----RNGWLRVYRDD-IPESTWISLWAMQDSNHVKFY-----ROAGADNLNPNYFT	858
Qy	1519	-----FELCLPDGDV-----SIHGAVGASELANTAKAEADAADAK	1553
Db	859	IDLNKEQWILRTQGEVAIP-GDVWRNVLTIKLQTRSVH-----IVAPTPKS--AERWAA	909
Qy	1554	LLGNS-----LLKLEGDDRL-DMNCITLPPS-----DQVVLVGTREGVYALANVL	1595
Db	910	CLQNAQTRRMKNKSPSIAEYSCULTJUSLPNNLKIFKAHTIEDWILFATQGTGFFTSIS	969
Qy	1596	K-NSLTHVPDIGAIVFYI1KDEKLKIMTAGBERALCLVDVKVKVQSQAQSHLPAQPDIS	1654
Db	970	QPRNPTRIAGPSVTSLEVMSEINCAMITNSNRQLALIPLDSLTLMQSTH-----PSIR	1025
Qy	1655	PN1FEAVKGCHLPGAGK1ENGLIC1CAAMPSKVILRYNENISKYICIRKEIETSEPCSCIH	1714
Db	1026	AEVLPEFGVHTIRXHQDQOGRFLVSDSDTQJHRIKYNSTRDVFHAFKLVVPEPVSTIE	1085
Qy	1715	FTNYSILIGTNKFEYEDMKQYTLFEEFLDNHDSHAPAVFAAS-----SNSFPVSIVQVNS	1769
Db	1086	STAGIFASDIFYVYPL-----DHQSPSDVSARKLMPRRSDYPVSAHAITA	1133
Qy	1770	AGOREEYLLCFHEFGVDSYGRSRTDCLKSRPLAPAYAREPYLFVTHFNSLEVIEIQ	1829
Db	11134	-----NEILLAYQNHGIFVNLYGEOQRNQ11IEWEKMPIEFAYTSPULYIVHDDSEIIVQIS	1189
Qy	1830	ARSSAGTPARAYLD1ENPRYLOPA1SSGAIY1ASSYQDKLR1VICCKGNLKVESG-----	1883
Db	1190	K-----SSKETVLA-----EREVIACVNAHIVQSDGLIVSVS	1221
Qy	1884	-----THHRGPSTSRSPNKR	1900
Db	1222	SKDITEVHRFSA1CRGTGTR	1242

Search completed: July 3, 2004, 10:15:38  
Job time : 44 secs

C:Genetics:		Query Match		6.7%; Score 706.5; DB 2; Length 1256;	
A:Gene: CESP:W0289.2		Best Local Similarity		21.1%; Pred. No. 7.1e-13;	
A:Map position: 2		Matches		308; Conservative 292; Mismatches 528; Indels 333; Gaps 54;	
A:Introns: 27/3; 327/3; 670/3; 949/3; 1073/3					
Qy	545	QSRKIQIKQEQYQAOQVEEMRLMNQLEEDLVGARRSDLYSELSERSLAA-EFPRKZA	603		
Db	10	QSGKTPKSGKARYSSPYQVIDL--NNESQEDLTKRLLEAENIIQDLASERDALHESLVDKA	68		
Qy	604	TECHKLLKAKDGKPGVGVAKLEKINASQQLKIQELQELKSAKAKERAERELEKLQNR	663		
Db	69	GLNESVIERSNKVSTQETRIYRDVTLLEDLQKESQIRILQNRCLRLETEKQKQV--	126		
Qy	664	EDSSEGIRKKLVPAEERRHSLNKKVKRLTWERRENRLKDDIOTKSOQI----	QOWADKI	719	
Db	127	-DTISGYQEDLKENE-----IRIENLNSRLHLEDELSAKTHEIFSIGEELKNKT	175		
Qy	720	LELEEKHRE-----AOVSAOHLEVHLKQKHVEEKIKVLQNOIKKDLADKETLNNMOR	774		
Db	176	MKLNKNSQQTKLAEISSNR--NLBRKVQKFEELIVKD-----	QRSLEV	220	
Qy	775	HEBEAHKGIKLSQKAMINAMD--SKIRSLQRIVELSEPAKLAANSUFTQNNKAAQE	832		
Db	221	HQDQNTQ-KVLKEVKQLSDRLDYLTPKRDVSRKERDD-----	FLQFSAKIIE	269	
Qy	833	EMTSELFCQKFXYLETOAGKLEAQNKLFEQLEKISHQDHSKDKRLLELETRLEVSLEHE	892		
Db	270	ETWSELKDKNARLERUSE-----KEELVKVT-----	KEELQELQTVTQAMGDSE	315	
Qy	893	E-----QKLELRQUTEQLSLQERESQITLQAARAALESQLRQAKTELEETT--	A	942	
Db	316	QATKYLHAENNKLTQKADIRCDLLEARNLKGFDKRELEKQDEDAEDVRRIETLKK	375		
Qy	943	EAESEIQAHTARDEIQKPDALNSC---TVITDLEEQNLQITEDNAELNNQNFVLSK	998		
Db	376	NVEIELSKLLAERBEQIDELKSRVAGYEVUURDHEAVKNEAKAEKLNKNGAHL--	433		
Qy	999	QLDEASGANDIEIVLRSEVDHLREITEREMQLTSQKOTMEALKTTCMLEEQMDLEAL	1058		
Db	434	-----VMADQKSSHFK--TLKETAEGSRRRRAIE---QCNEVMVARIGLEA-	473		
Qy	1059	NDELLEKRWQEAWSVLGDEKSFQRCRVELQEMLDTQKSRADQRITERSQVVELA	1118		
Db	474	--SLENQKRYE-----QELE-MVKAENVQRAKTEFVKETIQETHLD	512		
Qy	1119	VKEHKAETILQQAQLEKQKAEKSLSD-KLNDLEKHAMLEMAKRSQQKLETERELKQR	1177		
Db	513	YREBELSKL----AKGGSHEADQSQDSLSRAKTTIQEVKADNKKLQQLILEVRQNSK	567		
Qy	1178	LLEBQAKLQOQMDLQKHIIFFLTQGLQGLADRADLLKT-----	ERSDLEYQLENIQ	1228	
Db	568	VLEENVKLR-----KGMABAEKIEBEFKRNWHSREAGERIQLCAKENEK	613		
Qy	1229	VLYSHEKVMKMG-TISQOTKLIQLQAKMDQPAKKKKGLFSRRKEDPALTPQVPLQVNE	1287		
Db	614	VLKVEEBELQERLEVEKEELVNYLQSQINTKQTKQPKL--SRSLTMSLT-SEV----	DTS	668	
Qy	1288	KLALEKKAQCAELBEALQKTRIELRSABEAAHRKATDHPHPSPTPATARQQIAMSIVR	1347		
Db	669	TYRVEVSEVRA--LEEQREELQATYLAERKKWADLQKSRSTANTTTLTS-----	TT	717	
Qy	1348	SPEHQPAMSLLPAPSSRRKESSTPEFSRRLKERMEHNTPHRPNVGLNMR-----	ATKC	1402	
Db	718	ATEVYSKACELSYSPAT-----MRHDIPKW--LEFRHYSVLWSMKC	756		
Qy	1403	AVCLDTHVHFGQAQKCLCEQVMCHPKGKSTCLPATCGLPAEYATHFTFAFRDKMNSGLQ	1462		
Db	757	SLCFVGITFTFYANKKCHCDVRVHASCAPRVNNTCGIPACQATYYQDA--QAVGGIG--	811		

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:17:35 ; Search time 105 Seconds  
(without alignments)  
6086.340 Million cell updates/sec

Title: US-10-017-216-2  
Perfect score: 10490  
Sequence: 1 MLKFKYGARNPLDAGAEPI.....QLNGEIROQVEKSVLRITDYC 2053

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA.\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US50\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US50\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10490	100.0	2053	13	US-10-017-216-2
2	10490	100.0	2053	14	US-10-325-430-12
3	10032.5	95.6	2054	13	US-10-415-011-21
4	10022.5	95.5	2054	13	US-10-028-946-2
5	10007	95.4	2066	12	US-09-964-956-9
6	10007	95.4	2066	12	US-10-262-511-14
7	10005	95.4	2053	12	US-09-964-956-11
8	10005	95.4	2053	12	US-10-262-511-2
9	9656	92.0	2055	13	US-10-017-216-4
10	9487.5	90.4	1958	13	US-10-028-946-4
11	7717.5	73.6	1841	12	US-09-964-956-40
12	7717.5	73.6	1841	13	US-10-017-216-5
13	7491.5	71.4	1597	12	US-09-964-956-41
14	7491.5	71.4	1597	13	US-10-017-216-6
15	6745.5	64.3	1441	15	US-10-412-897-3

16	6223.5	59.3	1286	12	US-09-964-956-38	Sequence 38, Appli
17	6223.5	59.3	1286	13	US-10-017-216-7	Sequence 7, Appli
18	4718.5	45.0	999	12	US-10-276-774-1487	Sequence 1487, Ap
19	4519.5	43.1	940	12	US-09-964-956-39	Sequence 39, Appl
20	4249.5	40.5	883	12	US-09-964-956-37	Sequence 37, Appl
21	4101	39.1	847	12	US-10-262-511-8	Sequence 8, Appli
22	4005.5	38.2	832	12	US-10-262-511-6	Sequence 6, Appli
23	3155	30.1	623	12	US-10-262-511-4	Sequence 4, Appli
24	2910	27.7	623	12	US-10-262-511-10	Sequence 10, Appl
25	2892.5	27.6	638	12	US-10-262-511-12	Sequence 12, Appl
26	2430	23.2	497	9	US-09-804-471A-2	Sequence 2, Appli
27	2430	23.2	497	14	US-10-238-709-2	Sequence 2, Appli
28	2430	23.2	497	16	US-10-724-594-2	Sequence 2, Appli
29	2425	23.1	497	16	US-10-311-034-7	Sequence 7, Appli
30	2185.5	20.6	494	9	US-09-804-471A-4	Sequence 4, Appli
31	2185.5	20.6	494	14	US-10-238-709-4	Sequence 4, Appli
32	2185.5	20.6	494	16	US-10-724-594-4	Sequence 4, Appli
33	1700	16.2	373	15	US-10-120-835-42	Sequence 42, Appl
34	1645	15.7	319	15	US-10-412-897-2	Sequence 2, Appli
35	1446	13.8	1770	16	US-10-433-794-19	Sequence 19, Appl
36	1421.5	13.6	1719	12	US-10-362-892-2	Sequence 2, Appli
37	1421.5	13.6	1719	15	US-10-288-798-2	Sequence 2, Appli
38	1400	13.3	1738	15	US-10-210-130-100	Sequence 100, App
39	1399.5	13.3	1664	15	US-10-210-130-102	Sequence 102, App
40	1382.5	13.2	1711	9	US-09-771-161A-219	Sequence 219, App
41	1382.5	13.2	1711	9	US-09-771-161A-220	Sequence 220, App
42	1382.5	13.2	1711	12	US-10-399-225-10	Sequence 10, Appl
43	1277	12.2	257	14	US-10-282-048-2	Sequence 2, Appli
44	1258	12.0	1572	12	US-10-333-314-20	Sequence 20, Appl
45	1245	11.9	1546	15	US-10-369-493-6347	Sequence 6347, Ap

ALIGNMENTS

RESULT 1  
US-10-017-216-2  
; Sequence 2, Application US/10017216  
; Publication No. US20020160483A1  
; GENERAL INFORMATION:  
; APPLICANT: KAPILLER-LIBERMAN, Rosana  
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prote  
; TITLE OF INVENTION: Kinase and Uses Therefor  
; FILE OF INVENTION: 10147-5701  
; CURRENT APPLICATION NUMBER: US/10/017,216  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: US 60/242,429  
; PRIOR FILING DATE: 2000-10-23  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 2053  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-017-216-2

Query Match	100.0%;	Score 10490;	DB 13;	Length 2053;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2053;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MLKFKYGARNPLDAGAAEFIASRASRLNLFQGGKPPMTQQQMSPLSREGILDALFVLFE	60	
Db	1	MLKFKYGARNPLDAGAAEFIASRASRLNLFQGGKPPMTQQQMSPLSREGILDALFVLFE	60	
Qy	61	ECSPALMKIKHVSFNVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVVVREKATG	120	
Db	61	ECSPALMKIKHVSFNVRKYSDTIAELQELQPSAKDFEVRSLVGCGHFAEVVVREKATG	120	
Qy	121	DIYAMKMKKALLAQEQVSPFEERNILSRSTSPWIPQLQYAFQDKNHLYLMEEQPGG	180	
Db	121	DIYAMKMKKALLAQEQVSPFEERNILSRSTSPWIPQLQYAFQDKNHLYLMEEQPGG	180	
Qy	181	DLLSLNRYEDQDENLIQFYLAELILAVHSVHLMGYVHRDIPKPNILVDRGTGHIKLVDF	240	



```
; SEQ ID NO 12
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-325-430-12

Query Match      100.0%; Score 10490; DB 14; Length 2053;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDAGAAEPIASGRNLNLFPGKPPFTTQQQMSPLSREGILDALFVLFE 60
Db 1 MLKFKYGARNPLDAGAAEPIASGRNLNLFPGKPPFTTQQQMSPLSREGILDALFVLFE 60
QY 61 EGSQPALMKIKHVSFNFRKYSYDTIAELOELQPSAKDFEVRSLVGCCHFAEYQVVRKATG 120
Db 61 EGSQPALMKIKHVSFNFRKYSYDTIAELOELQPSAKDFEVRSLVGCCHFAEYQVVRKATG 120
QY 121 DIYAMKVMKKALLAQEQVSVFEEBENILSRSTSPWIPOLQYAFQDNHLYLMESYQPGG 180
Db 121 DIYAMKVMKKALLAQEQVSVFEEBENILSRSTSPWIPOLQYAFQDNHLYLMESYQPGG 180
QY 181 DLLSLNRYVEDQDENLQFYLAELILAVHSVHLNGYVHRDIKPNILVDRGTGHIKLVDF 240
Db 181 DLLSLNRYVEDQDENLQFYLAELILAVHSVHLNGYVHRDIKPNILVDRGTGHIKLVDF 240
QY 241 GSAAKVNSNMVNAKLPICPTDYPMAPEVLTVMNGDGKGTGYGLDCDWSVGVIAYEMIYGR 300
Db 241 GSAAKVNSNMVNAKLPICPTDYPMAPEVLTVMNGDGKGTGYGLDCDWSVGVIAYEMIYGR 300
QY 301 SPFAGTSARTNNIMNFQRFKFPDDPKVSDFLDLIQSLCGOKERLKEGLCCHPFF 360
Db 301 SPFAGTSARTNNIMNFQRFKFPDDPKVSDFLDLIQSLCGOKERLKEGLCCHPFF 360
QY 361 SKIDWNINRNSPPFPVPTLKSDDDTSNFDPEPKNSWSSPCQLSPGFSGEELPFVGFS 420
Db 361 SKIDWNINRNSPPFPVPTLKSDDDTSNFDPEPKNSWSSPCQLSPGFSGEELPFVGFS 420
QY 421 YSKALGILGRSSVSGLDSPAKTSMEKKLIKSKELQDSQKCHKMEQEMTLRHRRVS 480
Db 421 YSKALGILGRSSVSGLDSPAKTSMEKKLIKSKELQDSQKCHKMEQEMTLRHRRVS 480
QY 481 EVEAVLSQKEVELKASRTORSLEODLATYITECSSLSRSLRSEQARMESYQEDDKALQLLH 540
Db 481 EVEAVLSQKEVELKASRTORSLEODLATYITECSSLSRSLRSEQARMESYQEDDKALQLLH 540
QY 541 DIREQSRKLQETKEQEYQAEVEMLMNMNQLEEDLVARRRSDLYSELSRESRLAAEEFK 600
Db 541 DIREQSRKLQETKEQEYQAEVEMLMNMNQLEEDLVARRRSDLYSELSRESRLAAEEFK 600
QY 601 RKATECOHKLKAKQCGKPEVGEYAKLEKINAEQOLKIQELQEKLEKAAKRAERLEBKL 660
Db 601 RKATECOHKLKAKQCGKPEVGEYAKLEKINAEQOLKIQELQEKLEKAAKRAERLEBKL 660
QY 661 QNREDSSEGIKKLVEABERRHSLNKVKRLTMRERENRLKDDIQTQSQOIQQWADKIL 720
Db 661 QNREDSSEGIKKLVEABERRHSLNKVKRLTMRERENRLKDDIQTQSQOIQQWADKIL 720
QY 721 ELEKHREQAQSAHLEVLKQKQEHYEEKIKVLDNQIKKDLADKETLENMQRHEEBAH 780
Db 721 ELEKHREQAQSAHLEVLKQKQEHYEEKIKVLDNQIKKDLADKETLENMQRHEEBAH 780
QY 781 EKGKILSEOKAMINAWDSKIRSEORIVELSPANKLAANSISLFTORNKQAEEMISELRQ 840
Db 781 EKGKILSEOKAMINAWDSKIRSEORIVELSPANKLAANSISLFTORNKQAEEMISELRQ 840
QY 841 QKPYLETQAGKLEAQNRLKEOLEKISHQDSDKNRLLELETRLRREVSLHEHEQKLELR 900
Db 841 QKPYLETQAGKLEAQNRLKEOLEKISHQDSDKNRLLELETRLRREVSLHEHEQKLELR 900
QY 901 QUTELQLSQERESQITALQAARAALESQOLQAQKTELETTAAABEEIQALTARDEIQR 960
Db 901 QUTELQLSQERESQITALQAARAALESQOLQAQKTELETTAAABEEIQALTARDEIQR 960
```

```
QY 961 KFDALRNSCTVIITDLBEQNLQLTEDNAELNNQNFYLSKQLDEASGANDEIVQLRSEVDHL 1020
Db 961 KFDALRNSCTVIITDLBEQNLQLTEDNAELNNQNFYLSKQLDEASGANDEIVQLRSEVDHL 1020
QY 1021 RREITEREMQLTQKOTMEALKTTCTMLEBQVMDLEALNDELLEKEKQEWAWRSVLGDEK 1080
Db 1021 RREITEREMQLTQKOTMEALKTTCTMLEBQVMDLEALNDELLEKEKQEWAWRSVLGDEK 1080
QY 1081 SQFECRVRELQRLMDITBQKSRARADQRTESRQVVELAVKHEKAEILALQALKEQKLKA 1140
Db 1081 SQFECRVRELQRLMDITBQKSRARADQRTESRQVVELAVKHEKAEILALQALKEQKLKA 1140
QY 1141 ESLSDKLNDLEKXHAMLENNARSILOQKLETERELKORLLEEQAKLQOQMDLQKNHIFRLT 1200
Db 1141 ESLSDKLNDLEKXHAMLENNARSILOQKLETERELKORLLEEQAKLQOQMDLQKNHIFRLT 1200
QY 1201 QGLQALDRADLAKTERSLEYOLENIQVLYSHEKVMKEGTISQOTKILIDFLQAKMDQPA 1260
Db 1201 QGLQALDRADLAKTERSLEYOLENIQVLYSHEKVMKEGTISQOTKILIDFLQAKMDQPA 1260
QY 1261 KKKKGLFSRKEOPALPTQVPLQVNELKLALEKEKARCAELEALOKTRTELRSAREEA 1320
Db 1261 KKKKGLFSRKEOPALPTQVPLQVNELKLALEKEKARCAELEALOKTRTELRSAREEA 1320
QY 1321 HRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAPPSSRRKESSTPEEFSRRLK 1380
Db 1321 HRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAPPSSRRKESSTPEEFSRRLK 1380
QY 1381 ERMHNTPHFNVGLNNRATKCAVCLDTVHFGRQASKCLEQVMCHPKSTCJPATCGLP 1440
Db 1381 ERMHNTPHFNVGLNNRATKCAVCLDTVHFGRQASKCLEQVMCHPKSTCJPATCGLP 1440
QY 1441 AEYATHEATEAFCDKMSNPGLOTKEPSSSLHLEGMWKVPRNNKRGQOGMDRKVIVLEGSK 1500
Db 1441 AEYATHEATEAFCDKMSNPGLOTKEPSSSLHLEGMWKVPRNNKRGQOGMDRKVIVLEGSK 1500
QY 1501 VLIYDNEARAGORPVEEFELCLPDGQVSIHGAVGASELANATAKAEKAEADAKLGNLSL 1560
Db 1501 VLIYDNEARAGORPVEEFELCLPDGQVSIHGAVGASELANATAKAEKAEADAKLGNLSL 1560
QY 1561 KLGDDRLDNVCTLPFSDQVVLVGTBEGYALNVLKNSLTHVPGIGAVFOIYIIKOLEKL 1620
Db 1561 KLGDDRLDNVCTLPFSDQVVLVGTBEGYALNVLKNSLTHVPGIGAVFOIYIIKOLEKL 1620
QY 1621 LMTAGEERALCLVDVKKVQSLAQSHLPAQPDISPNIPEAVKCHLFGAGKIENGLCICA 1680
Db 1621 LMTAGEERALCLVDVKKVQSLAQSHLPAQPDISPNIPEAVKCHLFGAGKIENGLCICA 1680
QY 1681 AMPKSVVILRYNENLSKYCIKKEIETSEPCSCHTFTNYSLITCTNKPFYETDMKQYTLFEP 1740
Db 1681 AMPKSVVILRYNENLSKYCIKKEIETSEPCSCHTFTNYSLITCTNKPFYETDMKQYTLFEP 1740
QY 1741 LDKNDSLAPAVFAAASNSNPPVSIIVQNSAGQREBYLLCFHEFGVFVDSYGRSRTDDLK 1800
Db 1741 LDKNDSLAPAVFAAASNSNPPVSIIVQNSAGQREBYLLCFHEFGVFVDSYGRSRTDDLK 1800
QY 1801 WSRLLPAPAYREPYLVFVTHFNSLEVEIEIQARSSAGTPARAYLDIPNPRYLGPASSGAY 1860
Db 1801 WSRLLPAPAYREPYLVFVTHFNSLEVEIEIQARSSAGTPARAYLDIPNPRYLGPASSGAY 1860
QY 1861 LASSYQDKLRVICCKGNLVKESCTEHRHGPSTSRSSPNKRGPPPTYNEHTTKRVASSPAPP 1920
Db 1861 LASSYQDKLRVICCKGNLVKESCTEHRHGPSTSRSSPNKRGPPPTYNEHTTKRVASSPAPP 1920
QY 1921 EGSHPREPSTPHRYREGRTELARDKSPGRPLEREKSPGRMLSTRERSPGRLFEDSSRG 1980
Db 1921 EGSHPREPSTPHRYREGRTELARDKSPGRPLEREKSPGRMLSTRERSPGRLFEDSSRG 1980
QY 1981 RLPAGAVRTPLSQVNKRGQASQVTVNTVTVYDANKKLDNLIPANWSVLRIIQLNGEIR 2040
Db 1981 RLPAGAVRTPLSQVNKRGQASQVTVNTVTVYDANKKLDNLIPANWSVLRIIQLNGEIR 2040
```

QY	2041	QVQKSVLRDTC	2053	1	MLFKYGCARNPLDAGAEPIASRASRLNLFPOGKPPMTQQQMSPLSREGILDALFVLPE	60
Db	2041	QVQKSVLRDTC	2053	1	MLFKYGCARNPLDAGAEPIASRASRLNLFPOGKPPMTQQQMSPLSREGILDALFVLPE	60
RESULT 3						
US-10-415-011-21						
; Sequence 21, Application US/10415011						
; Publication No. US20040053394A1						
; GENERAL INFORMATION:						
; APPLICANT: INCYTE CORPORATION						
; APPLICANT: GURUPAJAN, Rajagopal						
; APPLICANT: BAUGHN, Mariah R.						
; APPLICANT: CHAWLA, Narinder K.						
; APPLICANT: ELLIOTT, Vicki S.						
; APPLICANT: XU, Yuming						
; APPLICANT: ARVIZU, Chandra S.						
; APPLICANT: YAO, Monique G.						
; APPLICANT: RAMKUMAR, Jayalaxmi						
; APPLICANT: DING, Li						
; APPLICANT: TANG, Y. Tom						
; APPLICANT: HAFALIA, April J.A.						
; APPLICANT: NGUYEN, Dannel B.						
; APPLICANT: GANDHI, Ameena R.						
; APPLICANT: LU, Yan						
; APPLICANT: YUE, Henry						
; APPLICANT: BURFORD, Neil						
; APPLICANT: BANDMAN, Olga						
; APPLICANT: TRIBOULEY, Catherine M.						
; APPLICANT: LAL, Preeti G.						
; APPLICANT: RECIPON, Shirley A.						
; APPLICANT: LU, Dyung Aina M.						
; APPLICANT: BOROMSKY, Mark L.						
; APPLICANT: THORNTON, Michael B.						
; APPLICANT: SWARNAKER Anita						
; APPLICANT: THANGAVELU, Kavitha						
; APPLICANT: KHAN, Farrah A.						
; APPLICANT: ISON, Craig H.						
; TITLE OF INVENTION: HUMAN KINASES						
; FILE REFERENCE: PI-0262 USN						
; CURRENT APPLICATION NUMBER: US/10/415,011						
; CURRENT FILING DATE: 2003-04-18						
; PRIOR APPLICATION NUMBER: PCT/US01/47728						
; PRIOR FILING DATE: 2001-10-20						
; PRIOR APPLICATION NUMBER: US 60/242,410						
; PRIOR FILING DATE: 2000-10-20						
; PRIOR APPLICATION NUMBER: US 60/244,068						
; PRIOR FILING DATE: 2000-10-27						
; PRIOR APPLICATION NUMBER: US 60/245,708						
; PRIOR FILING DATE: 2000-11-03						
; PRIOR APPLICATION NUMBER: US 60/247,672						
; PRIOR FILING DATE: 2000-11-09						
; PRIOR APPLICATION NUMBER: US 60/249,565						
; PRIOR FILING DATE: 2000-11-16						
; PRIOR APPLICATION NUMBER: US 60/252,730						
; PRIOR FILING DATE: 2000-11-22						
; PRIOR APPLICATION NUMBER: US 60/250,807						
; PRIOR FILING DATE: 2000-12-01						
; NUMBER OF SEQ ID NOS: 44						
; SOFTWARE: PERL Program						
; SEQ ID NO 21						
; LENGTH: 2054						
; TYPE: PRT						
; ORGANISM: Homo sapiens						
; FEATURE:						
; NAME/KEY: misc feature						
; OTHER INFORMATION: Incyte ID No. US20040053394A1 7484498CD1						
US-10-415-011-21						
Query Match 95.6%; Score 10032.5; DB 12; Length 2054;						
Best Local Similarity 95.8%; Pred. No. 0;						
Matches 1981; Conservative 2; Mismatches 4; Indels 81; Gaps 3;						

Db 1081 KEREWEASVLDGKESQCEVRRLQRMLDTEKQSRADQRITTESQVVELAYKEHA 1140  
QY 1125 EIALAQALKEQKLAESLSDKLNLEKHAMLENNARSLOQKLTTERELKORLEBOAK 1184  
Db 1141 EIALAQALKEQKLAESLSDKLNLEKHAMLENNARSLOQKLTTERELKORLEBOAK 1200  
QY 1185 LOQMDLQXNHIFRUTQGOEALDRADLLKTERSLEYOLENIQVLYSHEKVMGTTISQ 1244  
Db 1201 LOQMDLQXNHIFRUTQGOEALDRADLLKTERSLEYOLENIQVLYSHEKVMGTTISQ 1260  
QY 1245 QTKLIDFLOAKVDQPAKKKGLFSRKEDPALPTQVPLQYNELKLALEKEKARCAELBEA 1304  
Db 1261 QTKLIDFLOAKVDQPAKKK-----VPLQYNELKLALEKEKARCAELBEA 1305  
QY 1305 LOKTRIELRSAREEAAHRAKATDHPHPSTPATARQOIAMSAIVRSPEHOPSAMSLIAPSS 1364  
Db 1306 LOKTRIELRSAREEAAHRAKATDHPHPSTPATARQOIAMSAIVRSPEHOPSAMSLIAPSS 1365  
QY 1365 RRKESSTPEFSRRLKERHMHNIPIHFNVLGNLRATKCAVCLDTHVFGQASKCLECQVM 1424  
Db 1366 RRKESSTPEFSRRLKERHMHNIPIHFNVLGNLRATKCAVCLDTHVFGQASKCLECQVM 1425  
QY 1425 CHPKSTCLPATCGUPAEVATHTFAFCDDKXNSPGLQTKPESSSLHLEGWVKVPRNKR 1484  
Db 1426 CHPKSTCLPATCGUPAEVATHTFAFCDDKXNSPGLQTKPESSSLHLEGWVKVPRNKR 1485  
QY 1485 GQGGWDRKVIIVLEGGKVLIDYNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1544  
Db 1486 GQGGWDRKVIIVLEGGKVLIDYNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1545  
QY 1545 A-----EKARADAKL 1554  
Db 1546 ADVPYILKVESHPHTTCWFGRTYLLAPSPFPKQWVTALSVAGVRSREKAPADAKL 1605  
QY 1555 LGNSLLKLEGGDRLDNCTLPFSDQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFIYII 1614  
Db 1606 LGNSLLKLEGGDRLDNCTLPFSDQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFIYII 1665  
QY 1615 KOLEKLLMAGBERALCLVDKVKVQSLAQSHLPAQPOISPNIPEAVKGCHLFGAGKIEN 1674  
Db 1666 KOLEKLLMAGBERALCLVDKVKVQSLAQSHLPAQPOISPNIPEAVKGCHLFGAGKIEN 1725  
QY 1675 GLICIAAMPKSVILRYNENLSKYCIKRIEITSEPCSCIHFTNYSLIGTGNKFYIDMKQ 1734  
Db 1726 GLICIAAMPKSVILRYNENLSKYCIKRIEITSEPCSCIHFTNYSLIGTGNKFYIDMKQ 1785  
QY 1735 YTLBFLDKNDHSLAPAVFAASNSFPVSIQVNSAGQREBYLLCFHFEGFVDSYGRRS 1794  
Db 1786 YTLBFLDKNDHSLAPAVFAASNSFPVSIQVNSAGQREBYLLCFHFEGFVDSYGRRS 1845  
QY 1795 RTDDLKWSRLPLAFAYREPFLVTFHNSLEVIEIQARSAGTPARAYLDIPNPRYLGPAI 1854  
Db 1846 RTDDLKWSRLPLAFAYREPFLVTFHNSLEVIEIQARSAGTPARAYLDIPNPRYLGPAI 1905  
QY 1855 SSGAIVLASSYQDKLRVICCKNLVKESGTEHHRGPSTSRSSPNKRGPPTTNEHITKEVA 1914  
Db 1906 SSGAIVLASSYQDKLRVICCKNLVKESGTEHHRGPSTSRSSPNKRGPPTTNEHITKEVA 1965  
QY 1915 SSPAPPEGSPHPRBPTPHRYREGTELRDQSKPGRLEREKSPGRMLSTRRSPGSLF 1974  
Db 1966 SSPAPPEGSPHPRBPTPHRYREGTELRDQSKPGRLEREKSPGRMLSTRRSPGSLF 2025  
QY 1975 EDSSRRLPAGAVRTPLSQVKNKGQSA 2002  
Db 2026 EDSSRRLPAGAVRTPLSQVKNKWDQSS 2053

RESULT 4  
US-10-028-946-2  
; Sequence 2, Application US/10028946  
; Publication No. US20020123622A1  
; GENERAL INFORMATION:

; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Miranda, Maricar  
; APPLICANT: Friddle, Carl Johan  
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding t  
; FILE REFERENCE: LEX-0289-USA  
; CURRENT APPLICATION NUMBER: US/10/028,946  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: US 60/258,335  
; PRIOR FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 2054  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-028-946-2

Query Match 95.5%; Score 10022.5; DB 13; Length 2054;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 1979; Conservative 3; Mismatches 5; Indels 81; Gaps 3;  
QY 1 MLKFKYGARPLDAGAAEPIASRASRLNLFQGGPPMTQOQMSPLSREGILDALFVLFE 60  
Db 1 MLKFKYGARPLDAGAAEPIASRASRLNLFQGGPPMTQOQMSPLSREGILDALFVLFE 60  
QY 61 ECSOPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120  
Db 61 ECSOPALMKIKHVSNFVKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120  
QY 121 DIYAWKMKKALLAQOVSFEEERNILSRSTSPWIPQLQYAFQDKNHLVLYVEEYOPGG 180  
Db 121 DIYAWKMKKALLAQOVSFEEERNILSRSTSPWIPQLQYAFQDKNHLVLYVEEYOPGG 180  
QY 181 DLLSLNRYEDQLDENLIQFYLAELILAVHSVHLMGVVHRDIKPENILVDRTHGILKLVDF 240  
Db 181 DLLSLNRYEDQLDENLIQFYLAELILAVHSVHLMGVVHRDIKPENILVDRTHGILKLVDF 240  
QY 241 GSAAKMNSKNVAKLPIGTPDYWAPVLTWAGDGGKTYGLDCDWMSVGVIAEMIIYGR 300  
Db 241 GSAAKMNSKNVAKLPIGTPDYWAPVLTWAGDGGKTYGLDCDWMSVGVIAEMIIYGR 300  
QY 301 SPFAEGTSARTFNINMFQFLKFPDPPKVVSSDFLDLIQSLLCGQKERLKFEGLCCHPFF 360  
Db 301 SPFAEGTSARTFNINMFQFLKFPDPPKVVSSDFLDLIQSLLCGQKERLKFEGLCCHPFF 360  
QY 361 SKIDWNIRNSPPFPVTLKSDDDTSNFDPEKXNSWVSSPQCLSPSPGSGEELPFVGF 420  
Db 361 SKIDWNIRNSPPFPVTLKSDDDTSNFDPEKXNSWVSSPQCLSPSPGSGEELPFVGF 420  
QY 421 YSKALGILGRSESVVGLSDSPAKTSSMEKLLILKSKELQSDKCHKWEQMTLHRVS 480  
Db 421 YSKALGILGRSESVVGLSDSPAKTSSMEKLLILKSKELQSDKCHKWEQMTLHRVS 480  
QY 481 EVEAVLSQKEVELKASSETQSLLEQDLATYITECSSLKRSLERQARMEVSQEDDKALQLH 540  
Db 481 EVEAVLSQKEVELKASSETQSLLEQDLATYITECSSLKRSLERQARMEVSQEDDKALQLH 540  
QY 541 DIREQSKLQEIKEQYCAQVEEMRLMNOLEBDLYSARRSDLYSELSRLAAEFK 600  
Db 541 DIREQSKLQEIKEQYCAQVEEMRLMNOLEBDLYSARRSDLYSELSRLAAEFK 600  
QY 601 RKATECOQHLLKAKDQKQPEVGEYAKLEKINAEQQLKIQELQKLEKAKVAKTEATELLQ 648  
Db 601 RKATECOQHLLKAKDQKQPEVGEYAKLEKINAEQQLKIQELQKLEKAKVAKTEATELLQ 660  
QY 649 -----AKERARELEKLNREDSDSEGIRKKLVAEERPHSLKVKRLTWERENRLKDD 704  
Db 661 NIROAKERARELEKLNREDSDSEGIRKKLVAEERPHSLKVKRLTWERENRLKDD 720  
QY 705 IOTKSOQIQOMADKILEEESKHREAOVSAOHLVHLKQKQHOHYEEKIKVLJDNQIKKDLAD 764  
Db 721 IOTKSOQIQOMADKILEEESKHREAOVSAOHLVHLKQKQHOHYEEKIKVLJDNQIKKDLAD 780

QY 765 KETLENMORHEBEAHEKGIKILSEQAMINAMDSKIRSLERIVELSEANKLAANSSLET 824  
Db 781 KETLENMORHEBEAHEKGIKILSEQAMINAMDSKIRSLERIVELSEANKLAANSSLET 840  
QY 825 ORNKQAQBEIMSELROQKFYLETQAGKLEAQRKLEEQLEKISHQDSKDNRLLELETRL 884  
Db 841 ORNKQAQBEIMSELROQKFYLETQAGKLEAQRKLEEQLEKISHQDSKDNRLLELETRL 900  
QY 885 REVSLEHEBEOKLEIKQLTLEQLSLQERESQLTALQAAALLESQURQAkteLEETTABA 944  
Db 901 REVSLEHEBEOKLEIKQLTLEQLSLQERESQLTALQAAALLESQURQAkteLEETTABA 960  
QY 945 EBEIQALTARDEIQRKFDALRNSCTVITDLEBNQLNLTEDNAELNNQNFYLSKQLEDEAS 1004  
Db 961 EBEIQALTARDEIQRKFDALRNSCTVITDLEBNQLNLTEDNAELNNQNFYLSKQLEDEAS 1020  
QY 1005 GANDEIVQLRSEVDHLRREITEREMOLTSQKQTMELAKTTCTMLEBQVMDLEALANDELLE 1064  
Db 1021 GANDEIVQLRSEVDHLRREITEREMOLTSQKQTMELAKTTCTMLEBQVMDLEALANDELLE 1080  
QY 1065 KERQWEAARSVLGDEKQRECRVRELQRMLDTEKQSPARADQRIETESROVVELAVKEHKA 1124  
Db 1081 KERQWEAARSVLGDEKQRECRVRELQRMLDTEKQSPARADQRIETESROVVELAVKEHKA 1140  
QY 1125 EILALQALKEQKLEKAEISLSDKLNLDLEKKHAMELNARSLOQKLETERELKORLLEEQAK 1184  
Db 1141 EILALQALKEQKLEKAEISLSDKLNLDLEKKHAMELNARSLOQKLETERELKORLLEEQAK 1200  
QY 1185 LOQOMDLQKNIHFLRTQGLQOALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1244  
Db 1201 LOQOMDLQKNIHFLRTQGLQOALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1260  
QY 1245 QTKLIDFLOAKMDQPAKKKKGLFSRRKEDPALPTQVPLQYNELKLALEKAKARCALESEA 1304  
Db 1261 QTKLIDFLOAKMDQPAKKK-----VPLQYNELKLALEKAKARCALESEA 1305  
QY 1305 LQKTRIELSARBEAHRKATDHPHSTPATRQIQIAMSIAVRSBPBQPSAMSLAPSS 1364  
Db 1306 LQKTRIELSARBEAHRKATDHPHSTPATRQIQIAMSIAVRSBPBQPSAMSLAPSS 1365  
QY 1365 RKESSTPEFSRLKERVHHNIPHRFVNLNMRATKCAVCLDTHFGQASKCLECQVM 1424  
Db 1366 RKESSTPEFSRLKERVHHNIPHRFVNLNMRATKCAVCLDTHFGQASKCLECQVM 1425  
QY 1425 CHPKSTCLPATCGLPAEYATHFTEAFCDKMSPGLOTKEPSSSLHLEGMKVPNNKR 1484  
Db 1426 CHPKSTCLPATCGLPAEYATHFTEAFCDKMSPGLOTKEPSSSLHLEGMKVPNNKR 1485  
QY 1485 GQGWDRKVIIVLEGSKVLVYDNEAREAGORPVEEPFELCLPDGDVSIHGAAGSELANAK 1544  
Db 1486 GQGWDRKVIIVLEGSKVLVYDNEAREAGORPVEEPFELCLPDGDVSIHGAAGSELANAK 1545  
QY 1545 A-----EKASADAKL 1554  
Db 1546 ADVPYILKMESHPTTCPCRTLYLLA P S F PDKQRWVTAL E S V W A G R V S R E K A E A D A K L 1605  
QY 1555 LGSLLKLEGGDRDLNMCITLPSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVQIYII 1614  
Db 1606 LGSLLKLEGGDRDLNMCITLPSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVQIYII 1665  
QY 1615 KDLKLLMAGEBRALCLVDVKKVQSLPAQPDISPNI F A V K G C H L F G A G K I E N 1674  
Db 1666 KDLKLLMAGEBRALCLVDVKKVQSLPAQPDISPNI F A V K G C H L F G A G K I E N 1725  
QY 1675 GICICAAMPKVVILRYNENLSKYCIKKEIETSEPCSCIHFTNYSILIGTNKFPYIDMKQ 1734  
Db 1726 GICICAAMPKVVILRYNENLSKYCIKKEIETSEPCSCIHFTNYSILIGTNKFPYIDMKQ 1785  
QY 1735 YTLLEEFLLDKNDHSLAPAVFAAASNSPVSIVQVNSAGQREYLLCFHFEGFVDSYGRS 1794  
Db 1786 YTLLEEFLLDKNDHSLAPAVFAAASNSPVSIVQVNSAGQREYLLCFHFEGFVDSYGRS 1845  
QY 1795 RTDDLKWSRLPLAFAYREPYLFTVTHFNSLEVIEIQARSSAGTPARAYLDIPNRYLGPAL 1854

Db 1846 RTDDLKWSRLPLAFAYREPYLFTVTHFNSLEVIEIQARSSAGTPARAYLDIPNRYLGPAL 1905  
QY 1855 SSGAIYLAASSYQDKLRLVICCKGNLVKESGTEHHRGPTSSSPNKGPPTYNEHITKRVA 1914  
Db 1906 SSGAIYLAASSYQDKLRLVICCKGNLVKESGTEHHRGPTSSSPNKGPPTYNEHITKRVA 1965  
QY 1915 SSPAPPGSPHPRPSTPHRYREGRTELRDKSPGRPLEREKSPGRMLSTRERSRGLF 1974  
Db 1966 SSPAPPGSPHPRPSTPHRYREGRTELRDKSPGRPLEREKSPGRMLSTRERSRGLF 2025  
QY 1975 EDSRGLPAGAVRTPLSQVNKGRGQSA 2002  
Db 2026 EDSRGLPAGAVRTPLSQVNKWDQSS 2053

## RESULT 5

US-09-964-956-9  
; Sequence 9, Application US/09964956  
; Publication No. US20040043926A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David  
; APPLICANT: Gunther, Erik  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Leach, Martin D  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20040043926A1e1 Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-124  
; CURRENT APPLICATION NUMBER: US/09/964, 956  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/235, 631  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235, 633  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235, 808  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236, 064  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236, 065  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236, 066  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236, 135  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/237, 434  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/238, 321  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238, 399  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238, 396  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/276, 667  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/294, 823  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: 60/304, 868  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 2066

;	TYPE: PRT									
	ORGANISM: Homo sapiens									
	US-09-964-956-9									
	Query Match									
	Best Local Similarity 95.4%; Score 10007; DB 12; Length 2066;									
	Matches 1979; Conservative 3; Mismatches 7; Indels 82; Gaps 4;									
Qy	1	MLKFKYGARNPIDAGAAEPIASRASRLNLF	FOGKPPFMTQQQMSPLSREGILDALFVLFE	60						
Db	1	MLKFKYGARNPIDAGAAEPIASRASRLNLF	FOGKPPFMTQQQMSPLSREGILDALFVLFE	60						
Qy	61	ECSQPALMKIKHVSFVFKYSTTIAELOELQPSAKDFEVRSLVGGCHFAEYVQVREKATG	120							
Db	61	ECSQPALMKIKHVSFVFKYSTTIAELOELQPSAKDFEVRSLVGGCHFAEYVQVREKATG	120							
Qy	121	DIYAMKWKKALLAQEQVSPFEEERNILSRSTSPWIPOLQYAFQDKNHLVIMBEYQPGG	180							
Db	121	DIYAMKWKKALLAQEQVSPFEEERNILSRSTSPWIPOLQYAFQDKNHLVIMBEYQPGG	180							
Qy	181	DLLSLNRYEDQDENLTOFYLAELILAVHSVHLNGYVHRDIPKENILVDRTHIKLVDF	240							
Db	181	DLLSLNRYEDQDENLTOFYLAELILAVHSVHLNGYVHRDIPKENILVDRTHIKLVDF	240							
Qy	241	GSAAKVNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKGTGGLDCDWWSGVVIAYEMIYGR	300							
Db	241	GSAAKVNSNK-VNAKLPIGTPDYMAPEVLTVMNGDGKGTGGLDCDWWSGVVIAYEMIYGR	299							
Qy	301	SPFAGTSGARTNNIMNFORFLKFPDDPKVSSDFLDLIQSLLCGQKBLKXFEGLCCHPPF	360							
Db	300	SPFAGTSGARTNNIMNFORFLKFPDDPKVSSDFLDLIQSLLCGQKBLKXFEGLCCHPPF	359							
Qy	361	SKIDWNIRNSPPFVPTLKSDDDTSNFDPEPKNSWSSPCQLSPGFSGEBELPFVGFS	420							
Db	360	SKIDWNIRNAPPPFVPTLKSDDDTSNFDPEPKNSWSSPCQLSPGFSGEBELPFVGFS	419							
Qy	421	YSKALGILGRSESVVSGLDSPAKTSMKPKLLIKSKE-QDSQDKHWEQEMTLHRRVS	480							
Db	420	YSKALGILGRSESVVSGLDSPAKTSMKPKLLIKSKE-LQDSQDKCHKWEQEMTLHRRVS	479							
Qy	481	EVEAVLSQKVELKASEQRSLSLEODLATYITECSSLKRSLSEQARMVYSQEDDKALQLLH	540							
Db	480	EVEAVLSQKVELKASEQRSLSLEODLATYITECSSLKRSLSEQARMVYSQEDDKALQLLH	539							
Qy	541	DIREQSRKLQEIKEQEQYQAEVEMLMNQLEEDLVSAARRSDLYESELRESRLAAEFK	600							
Db	540	DIREQSRKLQEIKEQEQYQAEVEMLMNQLEEDLVSAARRSDLYESELRESRLAAEFK	599							
Qy	601	RKATECOHKLKAKOQGPVEGEYAKLEKINAEQOLKIQEIQEKLEKA-----	648							
Db	600	RKATECOHKLKAKOQGPVEGEYAKLEKINAEQOLKIQEIQEKLEKAVKASTEATELLQ	659							
Qy	649	-----AKERAERELEKIQNRDSSGIRKKLVEABERRHSLENKVKRLTETMERENRLKDD	704							
Db	660	NIRQAKERAERELEKIQNRDSSGIRKKLVEABERRHSLENKVKRLTETMERENRLKDD	719							
Qy	705	IQTKSQOIQQWADKILEEKEHREAOVSAQHLVHLKQKEQHYEBKIKVLDNQIKKDLAD	764							
Db	720	IQTKSQOIQQWADKILEEKEHREAOVSAQHLVHLKQKEQHYEBKIKVLDNQIKKDLAD	779							
Qy	765	KETLENMQRHEEAEHKGKILSEQKAMINANDSKIRLSQRIVELSEANKLAANSLSFT	824							
Db	780	KETLENMQRHEEAEHKGKILSEQKAMINANDSKIRLSQRIVELSEANKLAANSLSFT	839							
Qy	825	QRNMKAQEBMISLQKQKPYLETQAGKLEAQNKRKLEEQLEKISHQDSDKNRLLELETRL	884							
Db	840	QRNMKAQEBMISLQKQKPYLETQAGKLEAQNKRKLEEQLEKISHQDSDKNRLLELETRL	899							
Qy	885	REVSLHEBEQKLEIKRQLTLOLSIQEBESQITAQAAARAALESQIROAKTELEBTTAAE	944							
Db	900	REVSLHEBEQKLEIKRQLTLOLSIQEBESQITAQAAARAALESQIROAKTELEBTTAAE	959							
Qy	945	EBEIQALTAHRDEIQKFDALRNSCTVITDLEEQNLQTNEDNAELNNQNFYLSQQLDEAS	1004							

Db	960	EBEIQALTAHRDEIQKFDALRNSCTVITDLEEQNLQTNEDNAELNNQNFYLSQQLDEAS	1019							
Qy	1005	GANDEIVQLRSEVDHLRREITEREMQLTSQKQWMEALKTTCTMLSEQVMDLEALNDELLE	1064							
Db	1020	GANDEIVQLRSEVDHLRREITEREMQLTSQKQWMEALKTTCTMLSEQVMDLEALNDELLE	1079							
Qy	1065	KERQWEAWRSVLGDEKSGQFECRVREIQRMLDTEKQSRARADQRTISRQVVEAVKEHKA	1124							
Db	1080	KERQWEAWRSVLGDEKSGQFECRVREIQRMLDTEKQSRARADQRTISRQVVEAVKEHKA	1139							
Qy	1125	EIALQALQAEQKLABSLDKLNDLEKKHAMLEMNARSQQQKLETERELKORLLBEQAK	1184							
Db	1140	EIALQALQAEQKLABSLDKLNDLEKKHAMLEMNARSQQQKLETERELKORLLBEQAK	1199							
Qy	1185	LQQOMDLQKNHIFRLTQGLQEALEADRLDLKTERSDLEYQLENIQVLYSHKVKVMEGTISQ	1244							
Db	1200	LQQOMDLQKNHIFRLTQGLQEALEADRLDLKTERSDLEYQLENIQVLYSHKVKVMEGTISQ	1259							
Qy	1245	QTKLIDFLOAKMCPAKKKKGLFSRRKEDPALFTQVPLQVYNELKLALEKAKCAELEEA	1304							
Db	1260	QTKLIDFLOAKMCPAKKKK-----VPLQVYNELKLALEKAKCAELEEA	1304							
Qy	1305	LQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAPPSS	1364							
Db	1305	LQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSALVRSPEHQPSAMSLAPPSS	1364							
Qy	1365	RRKESSTPEFSRRLKERMHNI PHRFNVGLNVRATKCAVCLDTVHFGRQASKLEQCVWM	1424							
Db	1365	RRKESSTPEFSRRLKERMHNI PHRFNVGLNVRATKCAVCLDTVHFGRQASKLEQCVWM	1424							
Qy	1425	CHPKCSTCLPATCGLPAEYATHTFACRDKNMSPGLQTKPESSSLHLEGMKVPNNKR	1484							
Db	1425	CHPKCSTCLPATCGLPAEYATHTFACRDKNMSPGLQTKPESSSLHLEGMKVPNNKR	1484							
Qy	1485	GOQGDWKYIVLPGSKVLIYDNEARAGORPVEEFELCLPDGDUVSIHGAVGASELANTAK	1544							
Db	1485	GOQGDWKYIVLPGSKVLIYDNEARAGORPVEEFELCLPDGDUVSIHGAVGASELANTAK	1544							
Qy	1545	A-----EKAEADAKL	1554							
Db	1545	ADVPYILKMESHPTTCWPORUTLYLAPSPDKQWWTALLESVVAGGRVSRREKAEADAKL	1604							
Qy	1555	LGNLSLLKLEGGDRDLNMNCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFQIYII	1614							
Db	1605	LGNLSLLKLEGGDRDLNMNCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFQIYII	1664							
Qy	1615	KDLEKLLMIAGEBERALCLVDVKVKQSLAQSHLPAQPDISPNI FEAVKGCCHLFGAGKIEN	1674							
Db	1665	KDLEKLLMIAGEBERALCLVDVKVKQSLAQSHLPAQPDISPNI FEAVKGCCHLFGAGKIEN	1724							
Qy	1675	GLCITCAAMPKSVILRYNENLSKYCIKTEITSEPCSCIHFTNYSILIGTNKFEYIDMKQ	1734							
Db	1725	GLCITCAAMPKSVILRYNENLSKYCIKTEITSEPCSCIHFTNYSILIGTNKFEYIDMKQ	1784							
Qy	1735	YTLSEFLLDKNDHSLAPAVFAASNSFPVSIQVNSAGQREYLLCFHFPQVFDVSYGRRS	1794							
Db	1785	YTLSEFLLDKNDHSLAPAVFAASNSFPVSIQVNSAGQREYLLCFHFPQVFDVSYGRRS	1844							
Qy	1795	RTDCLKWSRLPLAFAYREPVLFTVTHNSLVEIIEIQAESASAGTPARAYLDIPNPRYLGPAL	1854							
Db	1845	RTDCLKWSRLPLAFAYREPVLFTVTHNSLVEIIEIQAESASAGTPARAYLDIPNPRYLGPAL	1904							
Qy	1855	SSGAIVLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPTYNEHITKRVA	1914							
Db	1905	SSGAIVLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPTYNEHITKRVA	1964							
Qy	1915	SSPAPEGPSHPPEPSTPHRYREGTRELDRDKSPGRFLEREKSPGRMLSTRERSRGLF	1974							
Db	1965	SSPAPEGPSHPPEPSTPHRYREGTRELDRDKSPGRFLEREKSPGRMLSTRERSRGLF	2024							
Qy	1975	EDSSRGLRPAQAVRTPLSQVNKGRGQASQV	2005							

Db 2025 EDSRGRPLPAGAVRTPSLSQVNKVRQHSACV 2055

RESULT 6

US-10-262-511-14

; Sequence 14, Application US/10262511

; Publication No. US20040038223A1

; GENERAL INFORMATION:

; APPLICANT: Smithson, Glenda

; APPLICANT: Milliet, Isabelle

; APPLICANT: Peyman, John A.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Ju, Jingfang

; APPLICANT: Li, Li

; APPLICANT: Guo, Xiaojia (Sasha)

; APPLICANT: Patturajan, Meera

; APPLICANT: Spytak, Kimberly A.

; APPLICANT: Esinger, Shlomit R.

; APPLICANT: Ellerman, Karen

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Ort, Tatiana

; APPLICANT: Gorman, Linda

; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Anderson, David W.

; APPLICANT: Zhong, Wei

; APPLICANT: Catterton, Elina

; APPLICANT: Ji, Weizhen

; APPLICANT: Miller, Charles E.

; APPLICANT: Rastelli, Luca

; APPLICANT: Stone, David J.

; APPLICANT: Pena, Carol E. A.

; APPLICANT: Shenoy, Suresh G.

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Rothenberg, Mark E.

; APPLICANT: Leach, Martin D.

; APPLICANT: Adege, Michele L.

; APPLICANT: Berghs, Constance

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-462C

; CURRENT APPLICATION NUMBER: US/10/262,511

; CURRENT FILING DATE: 2003-05-28

; PRIOR APPLICATION NUMBER: 60/326,483

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: 60/373,815

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/381,642

; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029

; PRIOR FILING DATE: 2002-10-09

; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-05-16

; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/373,260

; PRIOR FILING DATE: 2002-04-17

; PRIOR APPLICATION NUMBER: 60/373,826

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,435

; PRIOR FILING DATE: 2001-10-05

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 439

; SOFTWARE: CuraSeqList version 0.1

; SEQ ID NO 14

; LENGTH: 2066

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-262-511-14

Query Match

Best Local Similarity 95.4%; Score 10007; DB 12; Length 2066;

Matches 1979; Conservative 3; Mismatches 7; Indels 82; Gaps 4;

1065 KERQWEARSVLGDEKSPQECVREIQRLQMLDTEKQSRARADQRIITESTQVVELAVKEHA 1124  
1080 KERQWEARSVLGDEKSPQECVREIQRLQMLDTEKQSRARADQRIITESTQVVELAVKEHA 1139  
1125 EIALAQALKEQKLKAEISDLNLEKHAEMNARSLOOKLETERELKORLLEBOAK 1184  
1140 EIALAQALKEQKLKAEISDLNLEKHAEMNARSLOOKLETERELKORLLEBOAK 1199  
1185 LQQQMDLQKNHIFRUTQGLQEAALDRADLKLKTERSDLEYQLENIQVLYSHEKVMGTTISQ 1244  
1200 LQQQMDLQKNHIFRUTQGLQEAALDRADLKLKTERSDLEYQLENIQVLYSHEKVMGTTISQ 1259  
1245 QTKLIDFLQAKWDQPAKKKGLFSRKEDPALPTQVPLQYNELKLALEKEKARCACAELEEA 1304  
1260 QTKLIDFLQAKWDQPAKKK-----VPLQYNELKLALEKEKARCACAELEEA 1304  
1305 LQKTRIELSAREEAAHRKATDPHPSTPATARQAIAMSAIVRSPEHQPSAMSLAPSS 1364  
1305 LQKTRIELSAREEAAHRKATDPHPSTPATARQAIAMSAIVRSPEHQPSAMSLAPSS 1364  
1365 RKKESTPEFRRRLKERHNNIPIHRFNVLGNMRAKCAVCLDTHFRQASKCJLECOVM 1424  
1365 RKKESTPEFRRRLKERHNNIPIHRFNVLGNMRAKCAVCLDTHFRQASKCJLECOVM 1424  
1425 CHPKSTCJLPAFCGJPAEYATHFTFAFCRDKVNSPGLQTKEPSLSHLLEGWVKVPRNKR 1484  
1425 CHPKSTCJLPAFCGJPAEYATHFTFAFCRDKVNSPGLQTKEPSLSHLLEGWVKVPRNKR 1484  
1485 GQGGWDRKVIIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDSVHGAAGASELANAK 1544  
1485 GQGGWDRKVIIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDSVHGAAGASELANAK 1544  
1545 A-----EKAADAKL 1554  
1545 ADVPILKVESHPHTTCWPGRTLYLLAPSPDKQWVTALESVAVAGRVSRKEAADAKL 1604  
1555 LQNSLLKLEGGDRDLDMNCTLPESDQVVLVGTBEGLYALNVLNLSLTHVPGIGAVFQIYII 1614  
1605 LQNSLLKLEGGDRDLDMNCTLPESDQVVLVGTBEGLYALNVLNLSLTHVPGIGAVFQIYII 1664  
1615 KOLEKLMTAGBERALCLVDVKKVQSLAQSHLPAQPDISNPIFAVKGCHLFGAGKIEN 1674  
1665 KOLEKLMTAGBERALCLVDVKKVQSLAQSHLPAQPDISNPIFAVKGCHLFGAGKIEN 1724  
1675 GLCICAAMPKVVILRYNENLSKYCIKKEIETSEPCSCIHFTNYSLIGITNKFYEIDMKQ 1734  
1725 GLCICAAMPKVVILRYNENLSKYCIKKEIETSEPCSCIHFTNYSLIGITNKFYEIDMKQ 1784  
1735 YTLBFLDKNDHSLAPAVFAASNSFPVSVIVQVNSAGQREBYLLCFHFFGVFVDSYGRRS 1794  
1785 YTLBFLDKNDHSLAPAVFAASNSFPVSVIVQVNSAGQREBYLLCFHFFGVFVDSYGRRS 1844  
1795 RTDDLKWSRLPLAFVREPLYFVTHFNSLEVTETQARSAGTPARAYLDIPNRYLGPAL 1854  
1845 RTDDLKWSRLPLAFVREPLYFVTHFNSLEVTETQARSAGTPARAYLDIPNRYLGPAL 1904  
1855 SSGAIYLASSYQDKLRVICCKGNLVKESGTEHHRGFSTSRSSPNKRGPPTTYNEHITKRAVA 1914  
1905 SSGAIYLASSYQDKLRVICCKGNLVKESGTEHHRGFSTSRSSPNKRGPPTTYNEHITKRAVA 1964  
1915 SGPAPPEGSHPREBPTPHRYREGTELRDRKSPORPLEREXSPGRMLSTRERSPGRLF 1974  
1965 SGPAPPEGSHPREBPTPHRYREGTELRDRKSPORPLEREXSPGRMLSTRERSPGRLF 2024  
1975 EDSRGRPLPAGAVRTPLSQVKNKGRGQASQV 2005  
2025 EDSRGRPLPAGAVRTPLSQVKNKGRHSEACV 2055

RESULT 7  
US-09-964-956-11  
; Sequence 11, Application US/09964956  
; Publication No. US20040043926A1

GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David  
; APPLICANT: Gunther, Erik  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Leach, Martin D  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20040043926A1e1 Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-124  
; CURRENT APPLICATION NUMBER: US/09/964,956  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/235,631  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235,633  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235,808  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,064  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,065  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,066  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,135  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/237,434  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/238,321  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,399  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,396  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/276,667  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/294,823  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: 60/304,868  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2053  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-964-956-11

Query Match 95.4%; Score 10005; DB 12; Length 2053;  
Best Local Similarity 95.6%; Pred. No. 0;  
Matches 1978; Conservative 3; Mismatches 5; Indels 82; Gaps 4;

QY 1 MLKFKYGARNPLDGAABEPIASRASRLNLFQCKPPPTMQQMSPLSRREGILDALFVLFE 60  
Db 1 MLKFKYGARNPLDGAABEPIASRASRLNLFQCKPPPTMQQMSPLSRREGILDALFVLFE 60  
QY 61 ECSOPALMKIKHVSNFVRKYSDTIAELQELQPSAKDPFVRSVLCGCHFAEVQVVRKATG 120  
Db 61 ECSOPALMKIKHVSNFVRKYSDTIAELQELQPSAKDPFVRSVLCGCHFAEVQVVRKATG 120  
QY 121 DIYAMKVMKKALLAQEQVSPFEERNILSRSTSPMTIQLQYAFQDKNHLVLMEEYOPGG 180  
Db 121 DIYAMKVMKKALLAQEQVSPFEERNILSRSTSPMTIQLQYAFQDKNHLVLMEEYOPGG 180

QY 181 DLLSLNRYEDQDLENLQFYLAELILAVSHVLMGYVHRDIKPENILVDRTHIKLVDF 240  
DB 181 DLLSLNRYEDQDLENLQFYLAELILAVSHVLMGYVHRDIKPENILVDRTHIKLVDF 240  
QY 241 GSAKNSNKNVNAKPIGTPDYMAPEVLTVMGDQKGTGYGLDCDWWSGVYAYEMIYGR 300  
DB 241 GSAKNSNKNVNAKPIGTPDYMAPEVLTVMGDQKGTGYGLDCDWWSGVYAYEMIYGR 299  
QY 301 SPAEGTSARTFNIMNFQFLKFPDDPKVSSDFDLIOSSLGCGQERLKFGLCCHPFF 360  
DB 300 SPAEGTSARTFNIMNFQFLKFPDDPKVSSDFDLIOSSLGCGQERLKFGLCCHPFF 359  
QY 361 SKIDWNIRNSPPFPVPTLTKSDDDTGNFDEPEKNNSVSSPCOLSPGSGBELPVGFS 420  
DB 360 SKIDWNIRNSPPFPVPTLTKSDDDTGNFDEPEKNNSVSSPCOLSPGSGBELPVGFS 419  
QY 421 YSVALGILGRSESVVSGLDSPAKTSMEKKLILKSKELQSDQKCHKMQEMLHRRVS 480  
DB 420 YSVALGILGRSESVVSGLDSPAKTSMEKKLILKSKELQSDQKCHKMQEMLHRRVS 479  
QY 481 EVBAVLSQKEVELKASQTSRSLLEQDLATVITBCSSLSKRSLEQARMEVSQEDDKALQLIH 540  
DB 480 EVBAVLSQKEVELKASQTSRSLLEQDLATVITBCSSLSKRSLEQARMEVSQEDDKALQLIH 539  
QY 541 DIREQSKLOEIKBOEYQAOQVEWRLMNLQLEBDLVSAARRSDLYESELRESRLAABEFK 600  
DB 540 DIREQSKLOEIKBOEYQAOQVEWRLMNLQLEBDLVSAARRSDLYESELRESRLAABEFK 599  
QY 601 RKATECOHLLKAKDOQKPEVGYAKLEKINAEQQLKIQELQEKLEKA----- 648  
DB 600 RKATECOHLLKAKDOQKPEVGYAKLEKINAEQQLKIQELQEKLEKAVKASTEATLQ 659  
QY 649 -----AKERABRELEKLNQREDSSEGIRKKLVEABERRHSLNKVKLETMERENRLKDD 704  
DB 660 NIRQAKERABRELEKLNQREDSSEGIRKKLVEABERRHSLNKVKLETMERENRLKDD 719  
QY 705 IQTKSQOIQOMADKILLEEKHREAQVSAQHLEVLHKOQHYEKKIYVLDNQIKKOLAD 764  
DB 720 IQTKSQOIQOMADKILLEEKHREAQVSAQHLEVLHKOQHYEKKIYVLDNQIKKOLAD 779  
QY 765 KETLENMQRHEBEAHEKGIUSEQKAMINAMDSKIRSLQRIVELSEANKLAANSUFT 824  
DB 780 KETLENMQRHEBEAHEKGIUSEQKAMINAMDSKIRSLQRIVELSEANKLAANSUFT 839  
QY 825 ORNKQAQEMISELROOKFVLETOAQKLEAQNKLBEQLEKISHQDHSKNRLLELETRL 884  
DB 840 ORNKQAQEMISELROOKFVLETOAQKLEAQNKLBEQLEKISHQDHSKNRLLELETRL 899  
QY 885 REVSLEHEEQLKRLQTELQLSLQERESQTLAQARAALESQLRQAKTELETTAEA 944  
DB 900 REVSLEHEEQLKRLQTELQLSLQERESQTLAQARAALESQLRQAKTELETTAEA 959  
QY 945 BEBIOALTARHDEIQKFPALRNSCTVIITDLEBQNLQTEDNAELNNQNFYLSKOLDEAS 1004  
DB 960 BEBIOALTARHDEIQKFPALRNSCTVIITDLEBQNLQTEDNAELNNQNFYLSKOLDEAS 1019  
QY 1005 GANDEIVQLRSEVDHLRREITEREMOLTSQKOTMEALKTTCTMLEBQVMDLEALNDELLE 1064  
DB 1020 GANDEIVQLRSEVDHLRREITEREMOLTSQKOTMEALKTTCTMLEBQVMDLEALNDELLE 1079  
QY 1065 KERQWEANRSLVGDSEKQFECRVREIQRLMDTEKQSRARADQIRITESROVVELAVKEHKA 1124  
DB 1080 KERQWEANRSLVGDSEKQFECRVREIQRLMDTEKQSRARADQIRITESROVVELAVKEHKA 1139  
QY 1125 EILALQOALKKEOKLKAESLSDKLNDEKKHMLNNARSLQOKLETERELKORLLEEQAK 1184  
DB 1140 EILALQOALKKEOKLKAESLSDKLNDEKKHMLNNARSLQOKLETERELKORLLEEQAK 1199  
QY 1185 LQOQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1244  
DB 1200 LQOQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1259  
QY 1245 QTKLIDFLQAKMDQPAKKKKFLFSRRKEDPALPTQVPLQYNELKLALEKEKARCAELEEA 1304

DB 1260 QTKLIDFLQAKMDQPAKKKK-----VPLQYNELKLALEKEKARCAELEEA 1304  
QY 1305 LOKTRIELRSAREBAHRKATDHPSTPATARQOIAMSAIVRSPHQPSPAMSLLAPSS 1364  
DB 1305 LOKTRIELRSAREBAHRKATDHPSTPATARQOIAMSAIVRSPHQPSPAMSLLAPSS 1364  
QY 1365 RRKESSTPEEFSSRLKERMHEHNIPHRENVGLNMRATKCAVCLDTVHFGRQASKCLEQVM 1424  
DB 1365 RRKESSTPEEFSSRLKERMHEHNIPHRENVGLNMRATKCAVCLDTVHFGRQASKCLEQVM 1424  
QY 1425 CHPKCSTCLPATCGLPAYATHTEAFCRDKMSPLQTKPESSSLHLEGWKMVPRNNKR 1484  
DB 1425 CHPKCSTCLPATCGLPAYATHTEAFCRDKMSPLQTKPESSSLHLEGWKMVPRNNKR 1484  
QY 1485 GQQGWDKRYIVLEGSKVLIYDNEAREAGQRPVEFEFCLCPDGDVSIHGAVGASELANTAK 1544  
DB 1485 GQQGWDKRYIVLEGSKVLIYDNEAREAGQRPVEFEFCLCPDGDVSIHGAVGASELANTAK 1544  
QY 1545 A-----EKAEDAKL 1554  
DB 1545 ADVPYILKSMESHPTTCWPGRTLLYLLAPSPDKQWVTALESVVAGGRVSRKAEADAKL 1604  
QY 1555 LGNSLLKLEGDDRLDMNCTLPFSQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYII 1614  
DB 1605 LGNSLLKLEGDDRLDMNCTLPFSQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYII 1664  
QY 1615 KDLEKLLMIAGEERALCLVDVKKVQSLAQSHLPAQDISPNIFEAVKGCFLFGAGKIEN 1674  
DB 1665 KDLEKLLMIAGEERALCLVDVKKVQSLAQSHLPAQDISPNIFEAVKGCFLFGAGKIEN 1724  
QY 1675 GLCICCAAMPKSVILTRYNENLSKYCIRKETSETPSCSIHFTNYSLIGTNKFEYEDMKQ 1734  
DB 1725 GLCICCAAMPKSVILTRYNENLSKYCIRKETSETPSCSIHFTNYSLIGTNKFEYEDMKQ 1784  
QY 1735 YTLBEFLDKNDHSLAPAVFAASNSFPVSIQVNSAGOREEYLLCFHFGVFDVSYGRS 1794  
DB 1785 YTLBEFLDKNDHSLAPAVFAASNSFPVSIQVNSAGOREEYLLCFHFGVFDVSYGRS 1844  
QY 1795 RTDDLKWSRLPLAFAYREPVLVTHFNSLEVIIEIQARSSAGTPARAYLDIPNRYLGPAL 1854  
DB 1845 RTDDLKWSRLPLAFAYREPVLVTHFNSLEVIIEIQARSSAGTPARAYLDIPNRYLGPAL 1904  
QY 1855 SSGAIYLASSYQDKRLRVICCKGNLVKESGTEHHRGFPSTSRSSPNKRGPTTYNEHITKRV 1914  
DB 1905 SSGAIYLASSYQDKRLRVICCKGNLVKESGTEHHRGFPSTSRSSPNKRGPTTYNEHITKRV 1964  
QY 1915 SSPAPPSPGSHPREPSTPHRYREGRTLRDQKFGPRLEREKSPGRMLSTRRERSRGLF 1974  
DB 1965 SSPAPPSPGSHPREPSTPHRYREGRTLRDQKFGPRLEREKSPGRMLSTRRERSRGLF 2024  
QY 1975 EDSRGRPLPAGAVRTPLSQVKNKRGQSA 2002  
DB 2025 EDSRGRPLPAGAVRTPLSQVKNKWDQSS 2052

## RESULT 8

US-10-262-511-2  
; Sequence 2, Application US/10262511  
; Publication No. US20040039223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly R.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.

```

; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zernusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Wei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shency, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-2

Query Match          95.4%; Score 10005; DB 12; Length 2053;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1978; Conservative 3; Mismatches 5; Indels 82; Gaps 4;

Qy      1 MLKFKYGARNPDLGAGAPRIASRSLNLFPGCKPPFMTQQMSPLSREGILDALFVLFE 60
Db      1 MLKFKYGARNPDLGAGAPRIASRSLNLFPGCKPPFMTQQMSPLSREGILDALFVLFE 60

Qy      61 ECSQPALMKIKHVSFVRKSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVREKATG 120
Db      61 ECSQPALMKIKHVSFVRKSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVREKATG 120

Qy      121 DIYAMKMKKALLAQEQVSFEERNILSRSTSPWIPOLQYAFODKKNHLYLMEBYQPGG 180
Db      121 DIYAMKMKKALLAQEQVSFEERNILSRSTSPWIPOLQYAFODKKNHLYLMEBYQPGG 180

Qy      181 DLLSLNRYEDQDENLIQFYLAELILAVHSLMGYVHRDIPKPNILVDRGTGHIKLVD 240
Db      181 DLLSLNRYEDQDENLIQFYLAELILAVHSLMGYVHRDIPKPNILVDRGTGHIKLVD 240

Qy      241 GSAAKXNSNMVNAKLPIGTPTDYNAPEVLTVNMGDKGTGYGLDCDWSVGVIAEMIIYGR 300
Db      241 GSAAKXNSNK-VNAKLPIGTPTDYNAPEVLTVNMGDKGTGYGLDCDWSVGVIAEMIIYGR 299

```

```

Qy      301 SPFAEGTSAETNNIMNFQRLKFPDDPKVSSDFLDLIQSLLCGQKRLKFEGLCCHPEF 360
Db      300 SPFAEGTSARTFNIMNFQRLKFPDDPKVSSDFLDLIQSLLCGQKRLKFEGLCCHPEF 359

Qy      361 SKIDWNIRNSPPFFVPTLKSDDDTSNFDPEPKNSWVSSSPCQLSPSGFSGEELPFVGF 420
Db      360 SKIDWNIRNAPPPFFVPTLKSDDDTSNFDPEPKNSWVSSSPCQLSPSGFSGEELPFVGF 419

Qy      421 YSKALGITGRSESVVSGLDSPAKTSSMEKKLLIKSKLEQSDQDKCHQMEQMTLRHRVS 480
Db      420 YSKALGITGRSESVVSGLDSPAKTSSMEKKLLIKSKLEQSDQDKCHQMEQMTLRHRVS 479

Qy      481 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSILKRSLEQARMEVSEDDKALQLLH 540
Db      480 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSILKRSLEQARMEVSEDDKALQLLH 539

Qy      541 DIREQSRKLOEIKQEYQAOVEEMRLMMNOLBEDLVSAARRSDLYSELSRLAABEFK 600
Db      540 DIREQSRKLOEIKQEYQAOVEEMRLMMNOLBEDLVSAARRSDLYSELSRLAABEFK 599

Qy      601 RKATECOHKLKAKDOQKPEVGYAKLEKINAEQQLKIQLQEKLEKA----- 648
Db      600 RKATECOHKLKAKDOQKPEVGYAKLEKINAEQQLKIQLQEKLEKAVKASTAETELLQ 659

Qy      649 ----AKERAERELEKLNREDSEGIKKLVAEERRHSLKVKLETKWERENRLKDD 704
Db      660 NIQAERAERELEKLNREDSEGIKKLVAEERRHSLKVKLETKWERENRLKDD 719

Qy      705 IOTKSQOIQQMADKILEBEKHEBAQVSAOHLEVHLKQEOHYEKKVLDNQIKKDLAD 764
Db      720 IOTKSQOIQQMADKILEBEKHEBAQVSAOHLEVHLKQEOHYEKKVLDNQIKKDLAD 779

Qy      765 KETLENMQRHEBEAHEKGIKILSEQKAMINAMSKIRSLRQIRIVELSEANKLAANSFLT 824
Db      780 KETLENMQRHEBEAHEKGIKILSEQKAMINAMSKIRSLRQIRIVELSEANKLAANSFLT 839

Qy      825 QRNKAQEMISELRQOKFYLETQAGKLEAQNRLKEBOLKISHQDHSQKRLLELETRL 884
Db      840 QRNKAQEMISELRQOKFYLETQAGKLEAQNRLKEBOLKISHQDHSQKRLLELETRL 899

Qy      885 REVSLEHEBKLELKROLTELQLSLOERESQTLQAARAALLESQLRQAATELEETAAE 944
Db      900 REVSLEHEBKLELKROLTELQLSLOERESQTLQAARAALLESQLRQAATELEETAAE 959

Qy      945 EBEIQALTARDEIQRFKDALRNSCTVITDLEEQNLQEDNAELNNQNFYLSKQLDEAS 1004
Db      960 EBEIQALTARDEIQRFKDALRNSCTVITDLEEQNLQEDNAELNNQNFYLSKQLDEAS 1019

Qy      1005 GANDEIVOLSEVDHLREITEREMQJTSOKOTWEALKTTCTMLEQVMDLEALNDELLE 1064
Db      1020 GANDEIVOLSEVDHLREITEREMQJTSOKOTWEALKTTCTMLEQVMDLEALNDELLE 1079

Qy      1065 KERQWEAWRSVLGDEKSQFECRVRELQRMJLDTKQSRARADQRTESQVVELAVKEHKA 1124
Db      1080 KERQWEAWRSVLGDEKSQFECRVRELQRMJLDTKQSRARADQRTESQVVELAVKEHKA 1139

Qy      1125 EILALQALKEQKLEKASLSKDLNDEKKGAMLEMANRSLOQKLETERELKORLLEQAK 1184
Db      1140 EILALQALKEQKLEKASLSKDLNDEKKGAMLEMANRSLOQKLETERELKORLLEQAK 1199

Qy      1185 LQQOMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYQLENIQVLYSHEKVKWEGTISQ 1244
Db      1200 LQQOMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYQLENIQVLYSHEKVKWEGTISQ 1259

Qy      1245 QTKLIDFLQAMQPAKKKGLFSRRKEDPALPTQVPLQYNELKLALEKPKARCAELEEA 1304
Db      1260 QTKLIDFLQAMQPAKKK-----VPLQYNELKLALEKPKARCAELEEA 1304

Qy      1305 LQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSIVRSPHQPASVILAPSS 1364
Db      1305 LQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSIVRSPHQPASVILAPSS 1364

```

QY 1365 RRKESSTPEEFSRLXERMHNTI:PHRFNVGLNMRATKCAVCLDTVHFGQASKCLEQVM 1424  
Db 1365 RRKESSTPEEFSRLXERMHNTI:PHRFNVGLNMRATKCAVCLDTVHFGQASKCLEQVM 1424  
QY 1425 CHPKCSTCLPATCGLPAEYATHTFAFCRDKNMSPGLQTKPESSSLHLEGMWVPRNNKR 1484  
Db 1425 CHPKCSTCLPATCGLPAEYATHTFAFCRDKNMSPGLQTKPESSSLHLEGMWVPRNNKR 1484  
QY 1485 GQGWRKYIIVLBSGKVLIVDNEARAGORPVPEFELCLPDGDVSIHGAVGASELANTAK 1544  
Db 1485 GQGWRKYIIVLBSGKVLIVDNEARAGORPVPEFELCLPDGDVSIHGAVGASELANTAK 1544  
QY 1545 A-----EKASADAKL 1554  
Db 1545 ADVPYILKMESHPTTCWPGORTYLLAPFPDQKRWVTALESVAVAGGRVSRKAZADAKL 1604  
QY 1555 LGNSLLKLEGGDLNMCNLPFDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFQIYII 1614  
Db 1605 LGNSLLKLEGGDLNMCNLPFDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFQIYII 1664  
QY 1615 KDIKLLMAGEBRALCLVDVKVQSLAQSHLPAQPDISPNIPEAVKCHLFGAGKLEN 1674  
Db 1665 KDIKLLMAGEBRALCLVDVKVQSLAQSHLPAQPDISPNIPEAVKCHLFGAGKLEN 1724  
QY 1675 GLGICAAKPSKVILRYNENLSKYCIKRIETSEPCSCIHFTNYSILIGTNKFIYIDMKQ 1734  
Db 1725 GLGICAAKPSKVILRYNENLSKYCIKRIETSEPCSCIHFTNYSILIGTNKFIYIDMKQ 1784  
QY 1735 YTLBEFLDKNDHSLAPAVPAASNSPVSIVQVNSAGQREBYLLCPHEFGVFDVSGRRS 1794  
Db 1785 YTLBEFLDKNDHSLAPAVPAASNSPVSIVQVNSAGQREBYLLCPHEFGVFDVSGRRS 1844  
QY 1795 RTDDLKWSRLPLAFAREPYLFTVTHNSLEVEIQARSAGTPARAYLDIPNPRYLGPAL 1854  
Db 1845 RTDDLKWSRLPLAFAREPYLFTVTHNSLEVEIQARSAGTPARAYLDIPNPRYLGPAL 1904  
QY 1855 SSGAIYLAASYQDKLRVICCKGNLVKESGTEHHRGSTSRSPNKRGPPTVNEHITKVA 1914  
Db 1905 SSGAIYLAASYQDKLRVICCKGNLVKESGTEHHRGSTSRSPNKRGPPTVNEHITKVA 1964  
QY 1915 SSPAPPEGSHPREPSTPHRYREGRTLELRDKSPGRPLEREXSPGMLSTRERSPGRLF 1974  
Db 1965 SSPAPPEGSHPREPSTPHRYREGRTLELRDKSPGRPLEREXSPGMLSTRERSPGRLF 2024  
QY 1975 EDSSRGLPAGAVRTPLSQVNGKRGQSA 2002  
Db 2025 EDSSRGLPAGAVRTPLSQVNGKWDQSS 2052

RESULT 9  
US-10-017-216-4  
; Sequence 4, Application US/10017216  
; Publication No. US20020160483A1  
; GENERAL INFORMATION:  
; APPLICANT: KAPILLER-LIBERMANN, Rosana  
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prod  
; FILE OF INVENTION: Kinase and Uses Therefor  
; FILE REFERENCE: 10147-5701  
; CURRENT APPLICATION NUMBER: US/10/017,216  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: US 60/242,429  
; PRIOR FILING DATE: 2000-10-23  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 2055  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-017-216-4

Query Match 92.0%; Score 9656; DB 13; Length 2055;  
Best Local Similarity 91.9%; Pred. No. 0;  
Matches 1903; Conservative 41; Mismatches 42; Indels 84; Gaps 5;

QY 1 MLKPKYGARPLDAGAAEPIASRASRLNLFPOGKPPFMTQQQSPISREGILDALFVLFE 60  
Db 1 MLKPKYGRNPPPEASASEPIASRASRLNLFPOGKPPMLTQQQSAISREGMLDALFALFE 60  
QY 61 BCSPALMKIKHNSFVYKYSDTIAELQELQPSAKOFEVRSVLGCGHFAEVSQVREKATG 120  
Db 61 BCSPALMKIKHNSFVYKYSDTIAELQELQPSAKOFEVRSVLGCGHFAEVSQVREKATG 120  
QY 121 DIYAMKMKKALLAQQVQSFEEERNILSRSTPMIPQLOQYAFQDNKHLYLMEETQPG 180  
Db 121 DVIAMKMKKALLAQQVQSFEEERNILSRSTPMIPQLOQYAFQDNKHLYLMEETQPG 180  
QY 181 DLLSLNARYEDQDENLIFOYLAELILAVSHVLMGVVHRDIKPNILVDRTGHILKLVDF 240  
Db 181 DFLSLNARYEDQDENLIFOYLAELILAVSHVLMGVVHRDIKPNILVDRTGHILKLVDF 240  
QY 241 GSAAKMSNKNVNAKLPDGTDMAPVLTVMGDKGTGTYGLDCDWMVSGVIAIYEMIGR 300  
Db 241 GSAAKMSNKN - VDAKLPDGTDMAPVLTVMGDKGTGTYGLDCDWMVSGVIAIYEMIGR 299  
QY 301 SPFACHTSARTFNINMFORFLKFPDDPKYSSDFLDLIQSLICGOKERLKFGLCHPFF 360  
Db 300 TPFTEGTSARTFNINMFORFLKFPDDPKYSSBLLDLQLLQVQERLKFGLCHPFF 359  
QY 361 SKIDWNIRNSPPFPVPTLAKSDDDTSNFDPEPEKNSWVSSSPCQLSPGSGEELPFVGF 420  
Db 360 ARTDWNIRNSPPFPVPTLAKSDDDTSNFDPEPEKNSWAFILCVPAEPLAFSGEELPFVGF 419  
QY 421 YSKALGILGHSVSVGLDSPAITSMEKKLLIKSELQDSQKCHMQEOMTRLHRRVS 480  
Db 420 YSKALGILGHSVSVGLDSPAITSMEKKLLIKSELQDSQKCHMQEOMTRLHRRVS 479  
QY 481 EBEAVLSQKEVELKASTQSLLEQDLATYITCSSLKRLEQARMEVSEDDKALQLLH 540  
Db 480 EBEAVLSQKEVELKASTQSLLEQDLATYITCSSLKRLEQARMEVSEDDKALQLLH 539  
QY 541 DIREQSKLGEIKQEYQAOVEMRLMNQLEBDLSARRRDLSELESRESLAEEFK 600  
Db 540 DIREQSKLGEIKQEYQAOVEMRLMNQLEBDLSARRRDLSELESRESLAEEFK 599  
QY 601 RKATECOHLLKAKDQKPEVGVAKLEKINAEQQLKIQELQEKLEKA----- 648  
Db 600 RANECQHLKAKDQKPEVGVAKLEKINAEQQLKIQELQEKLEKAVKASTENTELLQ 659  
QY 649 ---AKERAERELEKLNREDSSGIRKKLVEABERRHSLNKKVRLTWERENLKOD 704  
Db 660 NIRQAKERAERELEKLNREDSSGIRKKLVEABERRHSLNKKVRLTWERENLKOD 719  
QY 705 IQTKSQOIQOMADKILEBKHREAOVSAOHLVHLKQKEHYEEKIKVLNDQIKKDLAD 764  
Db 720 IQTKSEIQOMADKILEBKHREAOVSAOHLVHLKQKEHYEEKIKVLNDQIKKDLAD 779  
QY 765 KETLENMMQREBEAHEKGIKILSEQAMINAMDSKIRSEQRIVELSEANKLAANSSFLT 824  
Db 780 KESLENMMQREBEAHEKGIKILSEQAMINAMDSKIRSEQRIVELSEANKLAANSSFLT 839  
QY 825 QRNMKAQEMISELRQOKFYLETOAGKLEAQNKLSEOLEKISHQDHSKVRLELETRL 884  
Db 840 QRNMKAQEMISELRQOKFYLETOAGKLEAQNKLSEOLEKISHQDHSKVRLELETRL 899  
QY 885 REVSLEHEEQKLELKROLTELOLSLOERESQTLAQARAALSESQLOAKTELEETAE 944  
Db 900 REVSLEHEEQKLELKROLTELOLSLOERESQTLAQARAALSESQLOAKTELEETAE 959  
QY 945 EEEIQALTARDEIQRFKDALRNSCTVITDLBQLNQLTEDNAELNNQNYLSKQJDEAS 1004  
Db 960 EEEIQALTARDEIQRFKDALRNSCTVITDLBQLNQLTEDNAELNNQNYLSKQJDEAS 1019  
QY 1005 GANDEIVQLRSEVDHLRREITEREMOLTSQKQNEALKTCTMLEBQVMDLEALNDELLE 1064  
Db 1020 GANDEIVQLRSEVDHLRREITEREMOLTSQKQNEALKTCTMLEBQVMDLEALNDELLE 1079

1065 KERQWAMRSLVLDKSKQSECEVRBRLQRLMDTEKQSRARADQRIITQSRQVVELAVKEKA 1124  
Db  
1080 KERQWAMRSLVLDKSKQSECEVRBRLQRLMDTEKQSRARADQRIITQSRQVVELAVKEKA 1139  
Qy 1125 EILALQALKEQKLKAEISDLKNDLEKKHAMLENNARSLOQKLETERELKORLEEOAK 1184  
Db 1140 EILALQALKEQKLKAEISDLKNDLEKKHAMLENNARSLOQKLETERELKORLEEOAK 1199  
Qy 1185 LQQQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYQLENIQVLYSHEKYKMEGTISQ 1244  
Db 1200 LQQQMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYQLENIQVLYSHEKYKMEGTISQ 1259  
Qy 1245 QTKLIDFLOAKMDQPAKXKGLFSRKEDPALPTQVPIQYNELKLALEKEKARCALKEEA 1304  
Db 1260 QTKLIDFLOAKMDQPAKXKGLFSRKEDPALPTQVPIQYNELKLALEKEKARCALKEEA 1304  
Qy 1305 LOKTRIELRSABEAAHRKATDHPHSTPATARQOIAMSIVRSPEHOPPSAMSLAPSS 1364  
Db 1305 LOKTRIELRSABEAAHRKATDHPHSTPATARQOIAMSIVRSPEHOPPSAMSLAPSS 1364  
Qy 1365 RKESSTPEEFGRRLKERMHNIPIHRFNVGLNMRATKCAVCLDTVHFGQASKCLECOVM 1424  
Db 1365 RKESSTPEEFGRRLKERMHNIPIHRFNVGLNMRATKCAVCLDTVHFGQASKCLECOVM 1424  
Qy 1425 CHPKSTCLPATCGLPABEATHTFAFCRDKVNSPCLQTKPESSLSHLLEGWKMVPRNKR 1484  
Db 1425 CHPKSTCLPATCGLPABEATHTFAFCRDKVNSPCLQTKPESSLSHLLEGWKMVPRNKR 1484  
Qy 1485 GOQWDRKYIVLEGSKVLIDNEAREAGQRPVEEFELCLPDGDSVIGHAVGASELANPAK 1544  
Db 1485 GOQWDRKYIVLEGSKVLIDNEAREAGQRPVEEFELCLPDGDSVIGHAVGASELANPAK 1544  
Qy 1545 A-----EKAADAKL 1554  
Db 1545 ADVPIYLNKSHPTTCWFGRTRYLLAPSPQKQWVTALSVVAGVRSREKAEADAKL 1604  
Qy 1555 LGNSLLKLEGGDRDLDMNCTLPESDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFIYII 1614  
Db 1605 LGNSLLKLEGGDRDLDMNCTLPESDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFIYII 1664  
Qy 1615 KOLEKLLTAGBERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN 1674  
Db 1665 KOLEKLLTAGBERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKGCHLFGAGKIEN 1724  
Qy 1675 GLCICAAKPSKVILRYNENLSKYCIKREIETSEPCSCIHFTNYSILITGNKFYIDMKQ 1734  
Db 1725 SLICICAAKPSKVILRYNENLSKYCIKREIETSEPCSCIHFTNYSILITGNKFYIDMKQ 1784  
Qy 1735 YLLEEFLDXNDHSLAPAVFAASNSGFPVSIQVNSAGQREERYLLCFHBFQVFDVSYGRRS 1794  
Db 1785 YLLEEFLDXNDHSLAPAVFAASNSGFPVSIQVNSAGQREERYLLCFHBFQVFDVSYGRRS 1844  
Qy 1795 RTDDLKWSRLPLAFAYREBYLFTVTHNSLEVTEIQARSSAGTPAPAYLIDIPNRYLGPAL 1854  
Db 1845 RTDDLKWSRLPLAFAYREBYLFTVTHNSLEVTEIQARSSAGTPAPAYLIDIPNRYLGPAL 1904  
Qy 1855 SSGAIYLAASYQDKLTVICCKGNLVKESGTEHHRGPTSRSSPNKRGPTTYNEHITKVA 1914  
Db 1905 SSGAIYLAASYQDKLTVICCKGNLVKESGTEHHRGPTSRSSPNKRGPTTYNEHITKVA 1964  
Qy 1915 SSPAPPEGSHPREPSTPHRY--REGRTBLRDKSGPRLEKEKSGFGRMLSTRRERSPCR 1972  
Db 1965 SSPAPPEGSHPREPSTPHRYDREGRTBLRDKSGPRLEKEKSGFGRMLSTRRERSPCR 2024  
Qy 1973 LPESSSRGRLPAGAVRTPLSQVKNKRGQSA 2002  
Db 2025 LPESSSRGRLPAGAVRTPLSQVKNKWDQSS 2054

RESULT 10  
US-10-028-946-4  
; Sequence 4, Application US/10028946  
; Publication No. US20020123622A1

; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Miranda, Maricar  
; APPLICANT: Fiddle, Carl Johan  
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding t  
; FILE REFERENCE: LEX-0289-USA  
; CURRENT APPLICATION NUMBER: US/10/028,946  
; PRIOR FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: US 60/258,335  
; PRIOR FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 1958  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-028-946-4

Query Match 90.4%; Score 9487.5; DB 13; Length 1958;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 1877; Conservative 1; Mismatches 1; Indels 81; Gaps 3;  
Qy 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQKPKPEMTQQQMSPLSRREGILDALFVLFE 60  
Db 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQKPKPEMTQQQMSPLSRREGILDALFVLFE 60  
Qy 61 ECSOPALMKTKHVSNFVRKYSDTIAELOELQPSAKDPEVRSLYCGGHFAEVOVVRKATG 120  
Db 61 ECSOPALMKTKHVSNFVRKYSDTIAELOELQPSAKDPEVRSLYCGGHFAEVOVVRKATG 120  
Qy 121 DIYAMKMKKALLAQOQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYOPGG 180  
Db 121 DIYAMKMKKALLAQOQVSFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYOPGG 180  
Qy 181 DLLSLNRYEDQDENLIQFYLAELILAVHSVHLMGVVHRDIPENILVDRTHGKILVDF 240  
Db 181 DLLSLNRYEDQDENLIQFYLAELILAVHSVHLMGVVHRDIPENILVDRTHGKILVDF 240  
Qy 241 GSAAKMNSKNVNAKLPIGTDPYMAPEVLTVMGDGGKTYGLDCDWSVGVIAVEMYIGR 300  
Db 241 GSAAKMNSKNVNAKLPIGTDPYMAPEVLTVMGDGGKTYGLDCDWSVGVIAVEMYIGR 300  
Qy 301 SPFAEGTSARTFNNIMNFQRLKFPDPPKVSDFDLTIQSLCCQKQERLKFEGLCCHPFF 360  
Db 301 SPFAEGTSARTFNNIMNFQRLKFPDPPKVSDFDLTIQSLCCQKQERLKFEGLCCHPFF 360  
Qy 361 SKIDWNNIRSPPPFVFTLKSDDDTSNFDEPEKNSWSSPCQLSPSGFSGEELPFVGF 420  
Db 361 SKIDWNNIRSPPPFVFTLKSDDDTSNFDEPEKNSWSSPCQLSPSGFSGEELPFVGF 420  
Qy 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKWQEMTRLHRRVS 480  
Db 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKWQEMTRLHRRVS 480  
Qy 481 EVEAVLSQKEVELKASTQSLLEQDLATYITBCSSLKRSLEQARMEVSEDKALQLLH 540  
Db 481 EVEAVLSQKEVELKASTQSLLEQDLATYITBCSSLKRSLEQARMEVSEDKALQLLH 540  
Qy 541 DIREQSKLOEIKBOEQVQAQVEEMRLMNNQLEEDLVASARRSDLYSESELRESLAAEFK 600  
Db 541 DIREQSKLOEIKBOEQVQAQVEEMRLMNNQLEEDLVASARRSDLYSESELRESLAAEFK 600  
Qy 601 RKATECQKLLKAKDOQKPEVGEYAKLEKINAEQQLKIQELQEKLEKA----- 648  
Db 601 RKATECQKLLKAKDOQKPEVGEYAKLEKINAEQQLKIQELQEKLEKAVKASTEATELLQ 660  
Qy 649 ----AKERARELEKLQNRDSSGIRKLVABEERHSLKVKRLFTWERENRLKDD 704  
Db 661 NIROAKERARELEKLQNRDSSGIRKLVABEERHSLKVKRLFTWERENRLKDD 720  
Qy 705 IQTKSQIQWADKILELEKHKREAAQVSAQHLEVHLKQEKQHYEETKIVLDNQIKKLAD 764  
Db 721 IQTKSQIQWADKILELEKHKREAAQVSAQHLEVHLKQEKQHYEETKIVLDNQIKKLAD 780

QY 765 KETLENNWQREBEAHEKGIKLEQKAMINAMDSKIRSLERIVELSEANKLAANSSLT 824  
Db 781 KETLENNWQREBEAHEKGIKLEQKAMINAMDSKIRSLERIVELSEANKLAANSSLT 840  
QY 825 QRNKAAQEMISLROCKFYLETQAGKLEAQNPKLEBEQLEKISHQDHSKDKRILLELETRL 884  
Db 841 QRNKAAQEMISLROCKFYLETQAGKLEAQNPKLEBEQLEKISHQDHSKDKRILLELETRL 900  
QY 885 REVSLEHEEOKLEKQLELQLESLQERESQLTALQARAALLESQLEAKTELEETTAA 944  
Db 901 REVSLEHEEOKLEKQLELQLESLQERESQLTALQARAALLESQLEAKTELEETTAA 960  
QY 945 EEBIQALTARHDEIQRKFDALRNSCTVITDLEBQLNQLTEDNAELNQNPFYLSKQLEDEAS 1004  
Db 961 EEBIQALTARHDEIQRKFDALRNSCTVITDLEBQLNQLTEDNAELNQNPFYLSKQLEDEAS 1020  
QY 1005 GANDEIVOLRSEVDHRLREITEREMOLTSQKOTMEALKTTCTMLEEQVMDLEALNDELLE 1064  
Db 1021 GANDEIVOLRSEVDHRLREITEREMOLTSQKOTMEALKTTCTMLEEQVMDLEALNDELLE 1080  
QY 1065 KERQWEAWRSVLGDEKSQFCRVRLEQRMIDTEKQSRARADQRIITBSROVVELAVKEHKA 1124  
Db 1081 KERQWEAWRSVLGDEKSQFCRVRLEQRMIDTEKQSRARADQRIITBSROVVELAVKEHKA 1140  
QY 1125 EIILAOALKEQKLEKESLSDKLNDEKXHAMLENNARSLOOKLETERELKORLLEEQAK 1184  
Db 1141 EIILAOALKEQKLEKESLSDKLNDEKXHAMLENNARSLOOKLETERELKORLLEEQAK 1200  
QY 1185 LOQOMDLQKHIFRLTQGLQEQALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1244  
Db 1201 LOQOMDLQKHIFRLTQGLQEQALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1260  
QY 1245 QTKLIDFLQAKMDQPAKXKGLFSRRKEDPALPTQVPLQYNELKJALEKEKARCAELSEA 1304  
Db 1261 QTKLIDFLQAKMDQPAKXKGLFSRRKEDPALPTQVPLQYNELKJALEKEKARCAELSEA 1305  
QY 1305 LQKTRIELSAREAAHRKATDHPHSTPATARQQTAMSAIVRSBPHQPSAMSLAPSS 1364  
Db 1306 LQKTRIELSAREAAHRKATDHPHSTPATARQQTAMSAIVRSBPHQPSAMSLAPSS 1365  
QY 1365 RRKESSTPEFSRRLKERMHNIHPHFVNGLNMRATKCAVCLDTVHFGRQASKLECCQM 1424  
Db 1366 RRKESSTPEFSRRLKERMHNIHPHFVNGLNMRATKCAVCLDTVHFGRQASKLECCQM 1425  
QY 1425 CHPKCSCTLPATCGLPAEYATHTEAFCDKQNSPOLQTKPESSSLHLBGMWVFNKKR 1484  
Db 1426 CHPKCSCTLPATCGLPAEYATHTEAFCDKQNSPOLQTKPESSSLHLBGMWVFNKKR 1485  
QY 1485 GQGWDRKIVLGGSKVLIYDNEARAGORPVEEFELCLPDGDVSTHGAVGASELANTAK 1544  
Db 1486 GQGWDRKIVLGGSKVLIYDNEARAGORPVEEFELCLPDGDVSTHGAVGASELANTAK 1545  
QY 1545 A-----EKADADAKL 1554  
Db 1546 ADVPYILKMEHPHTTWPORLYLLAPSPDKQRWVTALSWAGGRVSRKAEADAKL 1605  
QY 1555 LGNSLLKLEGGDRDLNMCNTLPSSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVQIYII 1614  
Db 1606 LGNSLLKLEGGDRDLNMCNTLPSSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVQIYII 1665  
QY 1615 KOLEKLLMTAGEBALCLVDVKKVQSLAQSHLPAPQDIPSPNIFPAVKGCHLFGAGKTIEN 1674  
Db 1666 KOLEKLLMTAGEBALCLVDVKKVQSLAQSHLPAPQDIPSPNIFPAVKGCHLFGAGKTIEN 1725  
QY 1675 GLICICAMPSKVILRYNENLSKYCIKKEIETSEPCSCIHFTNYSLIGTGNKFYIDMKQ 1734  
Db 1726 GLICICAMPSKVILRYNENLSKYCIKKEIETSEPCSCIHFTNYSLIGTGNKFYIDMKQ 1785  
QY 1735 YTLLEFLDKNDHSLAPVAFPAASNSGFPVSIQVNSAGQREERYLLCFHFEGFVFDVSYGRS 1794  
Db 1786 YTLLEFLDKNDHSLAPVAFPAASNSGFPVSIQVNSAGQREERYLLCFHFEGFVFDVSYGRS 1845

QY 1795 RTDDLKWSRLPLAFAYREPYLFTVTHFNSLEVIEIQARSSAGTTPARAYLDIPNPRYLGPAL 1854  
Db 1846 RTDDLKWSRLPLAFAYREPYLFTVTHFNSLEVIEIQARSSAGTTPARAYLDIPNPRYLGPAL 1905  
QY 1855 SSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPSTSR 1894  
Db 1906 SSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPSTSR 1945  
RESULT 11  
US-09-964-956-40  
; Sequence 40, Application US/09964956  
; Publication No. US20040043926A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R  
; APPLICANT: Smithson, Glennnda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David  
; APPLICANT: Gunther, Erik  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Bugess, Catherine E  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Leach, Martin D  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20040043926A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-124  
; CURRENT APPLICATION NUMBER: US/09/964,956  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/235,631  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235,633  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235,808  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,064  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,065  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,066  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,135  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/237,434  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/238,321  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,399  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,396  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/276,667  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/294,823  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: 60/304,868  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 40  
; LENGTH: 1641  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-964-956-40  
Query Match 73.6%; Score 7717.5; DB 12; Length 1641;  
Best Local Similarity 90.8%; Pred. No. 0;  
Matches 1541; Conservative 15; Mismatches 16; Indels 125; Gaps 5;

QY 374 PFVPTLKSDDDTSNFDPEPKNSWVSSPCQLSPSPGSGBELPFVGFYSKALGILGRSES 433  
Db 1 PFVPTLKSDDDTSNFDPEPKNSWVSSVCQLSPSPGSGBELPFVGFYSKALGILGRSES 60  
QY 434 VVSGLDSPAKTSGMEKKLIIKSKELQSDQKCHKEQEMTRLHRRVSEVAVLSQKEVEL 493  
Db 61 VVSSLDSPAKVSSMEKKLIIKSKELQSDQKCHKEQEMTRLHRRVSEVAVLSQKEVEL 120  
QY 494 KASETORSLEODLATYITECSSLKRSLEQARMEVSEQDDKALQLLHDIREQSRKLQEI 553  
Db 121 KASETORSLEODLATYITECSSLKRSLEQARMEVSEQDDKALQLLHDIREQSRKLQEI 180  
QY 554 EBYQAQVEEMRLMNNQLEEDLVASARRSDLYESELRESRLAABPKKATCQCHKLKA 613  
Db 181 EBYQAQVEEMRLMNNQLEEDLVASARRSDLYESELRESRLAABPKKATCQCHKLKA 240  
QY 614 KQKGPEVEGEYAKLEKINAEQQLKIQELQEKLEKA-----AKERAEREL 657  
Db 241 KQKGPEVEGEYSKLEKINAEQQLKIQELQEKLEKAVKASTEATELLQIRQAKERAEREL 300  
QY 658 EKLNREDSEGIKKLVFAEERRHSLNKKVRLTETMERENRLKDDIQTKSQQIQQVAD 717  
Db 301 EKLNREDSEGIKKLVFAE----- 321  
QY 718 KILELEEKHREZAQVSAQHLVEHLKQEHYEBKIKVLDNOIKKOLADKETTLENNMORHEE 777  
Db 322 ---ELEEKHREZAQVSAQHLVEHLKQEHYEBKIKVLDNQIKKOLADKESLENNMORHEE 378  
QY 778 EAHKKGKILSEKAMINAMSKIRSLRIVELSEBANKLAANSIFTORNKAQEMISE 837  
Db 379 EAHKKGKILSEKAMINAMSKIRSLRIVELSEBANKLAANSIFTORNKAQEMISE 438  
QY 838 LRQKQFYLTOAGKLEAQNKLBEQLEKISHODSDQKRLLETRLETRLEVSLEHEEQKLE 897  
Db 439 LRQKQFYLTOAGKLEAQNKLBEQLEKISHODSDQKRLLETRLETRLEVSLEHEEQKLE 498  
QY 898 LRQKQFYLTOAGKLEAQNKLBEQLEKISHODSDQKRLLETRLETRLEVSLEHEEQKLE 957  
Db 499 LRQKQFYLTOAGKLEAQNKLBEQLEKISHODSDQKRLLETRLETRLEVSLEHEEQKLE 558  
QY 958 IQRKFDALNSCTVITDLEEQNLQNTEDNAELNNQNFYLSKQDDEASGANDIIVOLRSEV 1017  
Db 559 IQRKFDALNSCTVITDLEEQNLQNTEDNAELNNQNFYLSKQDDEASGANDIIVOLRSEV 618  
QY 1018 DHLRREITEREQLTQSKQTMELKTTCTMLBEQVMDLEALNDELLEKERQWEAERSVLG 1077  
Db 619 DHLRREITEREQLTQSKQTMELKTTCTMLBEQVMDLEALNDELLEKERQWEAERSVLG 678  
QY 1078 DEKSQPECVRELQRLMDTEKQSRADQRIITESTQVVELAVKEHKAIBILALQQALKKEQK 1137  
Db 679 DEKSQPECVRELQRLMDTEKQSRADQRIITESTQVVELAVKEHKAIBILALQQALKKEQK 738  
QY 1138 LKAESLSDKLNLEKKHAMLENKARSLOQKLETERELKORLLEEQAKLQQQMDLQKNHIF 1197  
Db 739 LKAESLSDKLNLEKKHAMLENKARSLOQKLETERELKORLLEEQAKLQQQMDLQKNHIF 798  
QY 1198 RLTQGLQELADRADLLKTERSLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKVD 1257  
Db 799 RLTQGLQELADRADLLKTERSLEYQLENIQVLYSHEKVMEGTISQQTKLIDFLQAKVD 858  
QY 1258 QPAKKKKG:FSRRKEDPALPTQVPOYNELKLALEKEKARCAREALQKTRIELRSARE 1317  
Db 859 QPAKKKKG:-----VPOYNELKLALEKEKARCAREALQKTRIELRSARE 903  
QY 1318 EAAHRKATDHPHPSTPATARQOIAMSAIVRSPHQPMSAMSLAPPSSRRKESSTPEEFSR 1377  
Db 904 EAAHRKATDHPHPSTPATARQOIAMSAIVRSPHQPMSAMSLAPPSSRRKESSTPEEFSR 963  
QY 1378 RLKERMHNIPIRNFVGLNMRATKCAVCLDTHVFRQASKCLEQCMCHPKGKSTCLPATC 1437  
Db 964 RLKERMHNIPIRNFVGLNMRATKCAVCLDTHVFRQASKCLEQCMCHPKGKSTCLPATC 1023

QY 1438 GLPAEYATHTEAFRCRDKNMNSPGLQTKPESSSLHLEGWKMVPRNNKQCGQWDRKYIVLE 1497  
Db 1024 GLPAEYATHTEAFRCRDKNMNSPGLQTKPESSSLHLEGWKMVPRNNKQCGQWDRKYIVLE 1083  
QY 1498 GSKVLIYDNEAREAGQRPVEEFELCLPDGQDVSIHGAVGASELANAKA----- 1545  
Db 1084 GSKVLIYDNEAREAGQRPVEEFELCLPDGQDVSIHGAVGASELANAKADVPYILKMESH 1143  
QY 1546 -----EKAADAKLGNLSLLLEGDDR 1567  
Db 1144 HTTCWPQRTLYLAPSPFDKQWWTALSVVAGVRVSRKAEADAKLGNLSLLLEGDDR 1203  
QY 1568 LDMNCTLIPFSDQVVLVGTTEGLYALNVLKNSLTHVPQIGAVFOIYIIKOLEKILMIAGEE 1627  
Db 1204 LDMNCTLIPFSDQVVLVGTTEGLYALNVLKNSLTHVPQIGAVFOIYIIKOLEKILMIAGEE 1263  
QY 1628 RALCLVDVKKVQKSLAQSHLPAQDISPNIFEAVKGCFLFGAGKIENGLCICAAMPKSV 1697  
Db 1264 RALCLVDVKKVQKSLAQSHLPAQDISPNIFEAVKGCFLFGAGKIENGLCICAAMPKSV 1323  
QY 1698 ILRNYENLSKYCIKKEIETSEPCSCIHFTNYSILIGNKFEYEDMKOYTLDEFLDKNDHS 1747  
Db 1324 ILRNYENLSKYCIKKEIETSEPCSCIHFTNYSILIGNKFEYEDMKOYTLDEFLDKNDHS 1393  
QY 1748 LAPAVFAASNSFPVSIQVNSAQOREEYLLCFHEFGVFDVSYGRRSRTDCLKWSRLPLA 1807  
Db 1384 LAPAVFAASNSFPVSIQVNSAQOREEYLLCFHEFGVFDVSYGRRSRTDCLKWSRLPLA 1443  
QY 1808 FAYREPVLFTVFNLSLEIVIOARSSAGTAPARAYLDIPNPRYLGPALSSGAIYLISSYQD 1867  
Db 1444 FAYREPVLFTVFNLSLEIVIOARSSAGTAPARAYLDIPNPRYLGPALSSGAIYLISSYQD 1503  
QY 1868 KLRVICCKGNLVKESGTEHHRGPTSRSSPNKGPPTTYNEHITKRVASSPAPPEGSPHPR 1927  
Db 1504 KLRVICCKGNLVKESGTEHHRGPTSRSSPNKGPPTTYNEHITKRVASSPAPPEGSPHPR 1563  
QY 1928 EPSTPHRY--REGTELRDKSGORPLERKSPGRLSTRRERSPGRLFDSSRGRILPAG 1985  
Db 1564 EPSTPHRYDRREGTELRDKSGORPLERKSPGRLSTRRERSPGRLFDSSRGRILPAG 1623  
QY 1986 AVRTPLSQVKNKGQSA 2002  
Db 1624 AVRTPLSQVKNKWDQSS 1640

RESULT 12  
US-10-017-216-5  
; Sequence 5, Application US/10017216  
; Publication No. US20020160483A1  
; GENERAL INFORMATION:  
; APPLICANT: KAPPELLER-LIBERMANN, Rosana  
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prote  
; FILE OF INVENTION: Kinase and Uses Therefor  
; FILE REFERENCE: 10147-5701  
; CURRENT APPLICATION NUMBER: US/10/017, 216  
; PRIOR FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: US 60/242, 429  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1641  
; TYPE: ERT  
; ORGANISM: Mus musculus  
US-10-017-216-5

Query Match 73.6%; Score 7717.5; DB 13; Length 1641;  
Best Local Similarity 90.8%; Pred. No. 0;  
Matches 1541; Conservative 15; Mismatches 16; Indels 125; Gaps 5;  
QY 374 PFVPTLKSDDDTSNFDPEPKNSWVSSPCQLSPSPGSGBELPFVGFYSKALGILGRSES 433  
Db 1 PFVPTLKSDDDTSNFDPEPKNSWVSSVCQLSPSPGSGBELPFVGFYSKALGILGRSES 60

QY 434 VVSGLDSPAKTSSMEKKLLIKSKELQSDQKCHKMEQEMTRLHRRVSEVAVLSQKEVEL 493  
DB 61 VVSGLDSPAKVSMMEKKLLIKSKELQSDQKCHKMEQENTRLHRRVSEVAVLSQKEVEL 120  
QY 494 KASETQBSLLEQDLATYITTECSSLSKESLEQARMEVSEQDDKALQLLHDIREQSRKLQELIK 553  
DB 121 KASETQBSLLEQDLATYITTECSSLSKESLEQARMEVSEQDDKALQLLHDIREQSRKLQELIK 180  
QY 554 EQSYQYQVEMRLMMNLQEDLVSARRSDLYSELSRESRLAAEPKPKATCQHKLKA 613  
DB 181 EQSYQYQVEMRLMMNLQEDLVSARRSDLYSELSRESRLAAEPKPKATCQHKLKA 240  
QY 614 KDQCKPEVEGYAKLEKINAEQQLKILOELQKLEKA-----AKERAEREL 657  
DB 241 KDQCKPEVEGYAKLEKINAEQQLKILOELQKLEKAVKASTEATELLOTRQAKERAEREL 300  
QY 658 EKQNRDSESGIRKKLVFAEERHSLNKVRLETMERRENRLKDDIQTKSQQIQOMAD 717  
DB 301 EKLNHREDSESGIRKKLVFAE----- 321  
QY 718 KILELEBKHEAQSQAHLVHLKQEQHYEKKIVLDNQIKKOLADKETLENMQRHEE 777  
DB 322 ---ELEBKHEAQSQAHLVHLKQEQHYEKKIVLDNQIKKOLADKESLENMQRHEE 378  
QY 778 EAHEKGIKILSEQKAMINAMDSKIRSLQRIVELSEANKLAANSFTQRMKAQEMISE 837  
DB 379 EAHEKGIKILSEQKAMINAMDSKIRSLQRIVELSEANKLAANSFTQRMKAQEMISE 438  
QY 838 LRQCKPYLETQAGKLEAQNKLBEQLEKISHQDSDXNRLLELETRLRVSLVSEHEQKLE 897  
DB 439 LRQCKPYLETQAGKLEAQNKLBEQLEKISHQDSDXNRLLELETRLRVSLVSEHEQKLE 498  
QY 898 LRQCKPYLETQAGKLEAQNKLBEQLEKISHQDSDXNRLLELETRLRVSLVSEHEQKLE 957  
DB 499 LRQCKPYLETQAGKLEAQNKLBEQLEKISHQDSDXNRLLELETRLRVSLVSEHEQKLE 558  
QY 958 IQKFPDALRNSCTVITDLEEQNLQEDNAELNNQNFYLSKQLEDEASGANDSIVQLRSEV 1017  
DB 559 IQKFPDALRNSCTVITDLEEQNLQEDNAELNNQNFYLSKQLEDEASGANDSIVQLRSEV 618  
QY 1018 DILRREITEREMOLTSQKOTMEALKTTCTMLBEQVMDLEALNDELLEKEROEAWRSVLG 1077  
DB 619 DILRREITEREMOLTSQKOTMEALKTTCTMLBEQVMDLEALNDELLEKEROEAWRSVLG 678  
QY 1078 DEKSQFECVAVRELOQMLDTEKQSRARADQRIETESQVVELAVKEHKAETLALQALKEQK 1137  
DB 679 DEKSQFECVAVRELOQMLDTEKQSRARADQRIETESQVVELAVKEHKAETLALQALKEQK 738  
QY 1138 LKAEISDKLNDLEKHAEMLENNARSLOQKLETERELKQLEAEOALQOQMDLQKNHIF 1197  
DB 739 LKAEISDKLNDLEKHAEMLENNARSLOQKLETERELKQLEAEOALQOQMDLQKNHIF 798  
QY 1198 RLQOGLQEAELDRADLLEKTERSLEYOLENIVLYSHKVKMEGTISQOQTKLIDFLQAKMD 1257  
DB 799 RLQOGLQEAELDRADLLEKTERSLEYOLENIVLYSHKVKMEGTISQOQTKLIDFLQAKMD 858  
QY 1258 QPAKKKKGLFSRKKEDPALPTQVPLQYNELKLALEKKEKARCALBEALQKTRIEERSARE 1317  
DB 859 QPAKKKK-----VPLQYNELKLALEKKEKARCALBEALQKTRIEERSARE 903  
QY 1318 EAAHRAKATDHPHSTPATARQOIAMSIVRSPEHQPMSAMSLAPPSSRRKESSTPEFSR 1377  
DB 904 EAAHRAKATDHPHSTPATARQOIAMSIVRSPEHQPMSAMSLAPPSSRRKESSTPEFSR 963  
QY 1378 RLKERMHNIPIHFNVLGNLRATKCAVCLDTVHFGQASKLEQCMCHPKSTCLPATC 1437  
DB 964 RLKERMHNIPIHFNVLGNLRATKCAVCLDTVHFGQASKLEQCMCHPKSTCLPATC 1023  
QY 1438 GLPAEYATHFTFAFCDDKXNSPGLQKPESSSLHLEGMKVPNNKRGQOQGHDRXIVILE 1497  
DB 1024 GLPAEYATHFTFAFCDDKXNSPGLQKPESSSLHLEGMKVPNNKRGQOQGHDRXIVILE 1083

QY 1498 GSKVLIYDNEAREAGORPVEEFELCLPDGDSVSHGAVGASELANATAK----- 1545  
DB 1084 GSKVLIYDNEAREAGORPVEEFELCLPDGDSVSHGAVGASELANATAKADVPYILKMESH 1143  
QY 1546 -----EKAEDAKLLGNSLLKLEGDDR 1567  
DB 1144 HTTCWPGRKTYLLAPSPFPKQRWVWTALESVVAGGRVSRKAEADAKLLGNSLLKLEGDDR 1203  
QY 1568 LDMNCTIPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYIILKLEKLLMAGEE 1627  
DB 1204 LDMNCTIPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYIILKLEKLLMAGEE 1263  
QY 1628 RALCLVDVKKVKQSLAQSHLPAQPDISPNIFFAVKGCCHLFGAGKIENGLCICAAMPKVV 1687  
DB 1264 RALCLVDVKKVKQSLAQSHLPAQPDISPNIFFAVKGCCHLFGAGKIENGLCICAAMPKVV 1323  
QY 1688 ILRYNENLSKYCIKRIKTEIETSEPCSIHFTNYSILIGTNKFIYEDMKQYTLLEEFLDKNDHS 1747  
DB 1324 ILRYNENLSKYCIKRIKTEIETSEPCSIHFTNYSILIGTNKFIYEDMKQYTLLEEFLDKNDHS 1383  
QY 1748 LAPAVFAASNSFPVSIQVNSAGOREYLLCFHEFGVFDVSYGRRSRTDCLKWSRLPLA 1807  
DB 1384 LAPAVFAASNSFPVSIQVNSAGOREYLLCFHEFGVFDVSYGRRSRTDCLKWSRLPLA 1443  
QY 1808 PAYREPLYFVTHFNLSLEVIQARSSAGTTPARAYLDTPNPRYLGPALSSGAIYASSYQD 1867  
DB 1444 PAYREPLYFVTHFNLSLEVIQARSSAGTTPARAYLDTPNPRYLGPALSSGAIYASSYQD 1503  
QY 1868 KLRVICCKGNLVKESGTEHHRGPGSTSPSSPKGPPPTNYNEHITKRVASSPAPPEGSHPR 1927  
DB 1504 KLRVICCKGNLVKESGTEHHRGPGSTSPSSPKGPPPTNYNEHITKRVASSPAPPEGSHPR 1563  
QY 1928 EPSTPHRY--REGTELRDRDKSPORPLERKSPGRMLSTRERSPGRLPFDSSRGRPLAG 1985  
DB 1564 EPSTPHRYDRREGTELRDRDKSPORPLERKSPGRMLSTRERSPGRLPFDSSRGRPLAG 1623  
QY 1986 AVRTPLSQVNVKGRQSA 2002  
DB 1624 AVRTPLSQVNVKGRQSS 1640

## RESULT 13

US-09-964-956-41  
; Sequence 41, Application US/09964956  
; Publication No. US20040043926A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David  
; APPLICANT: Gunther, Erik  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Leach, Martin D  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20040043926A1e1 Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-124  
; CURRENT APPLICATION NUMBER: US/09/964, 956  
; PRIOR APPLICATION NUMBER: 2001-09-26  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235, 631  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235, 633  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235, 808  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236, 064

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,321
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,399
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,396
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/276,667
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/294,823
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/304,868
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-964-956-41

Query Match
Best Local Similarity 71.4%; Score 7491.5; DB 12; Length 1597;
Matches 1492; Conservative 16; Mismatches 12; Indels 83; Gaps 4;

QY 468 MEQEMTRLHRRVSEVAVLSQSEVELKASSETQSRSLLEDQLATYITECSSLKRSLEQARME 527
DB 9 MEQEMTRLHRRVSEVAVLSQSEVELKASSETQSRSLLEDQLATYITECSSLKRSLEQARME 68
QY 528 VSOEDDKALQLLHDIREQSRKLOEIKQEQYQACVEMRLMNMQLBEDLVASRRRSDLYES 587
DB 69 VSOEDDKALQLLHDIREQSRKLOEIKQEQYQACVEMRLMNMQLBEDLVASRRRSDLYES 128
QY 588 ELRESRLAAEFKFRKATECOHKLKAKDOQGEVGEYAKLEKINAEQQLKIQELOEKLEK 647
DB 129 ELRESRLAAEFKFRKANEGCHKLKAKDOQGEVGEYSKLEKINAEQQLKIQELOEKLEK 188
QY 648 A-----AKERARELEKLNREDSESGIRKKLVAEARRHSLENKVRL 691
DB 189 AVKASTEATELLQNTIRQAERARELEKLNREDSESGIKKKLVAEARRHSLENKVRL 248
QY 692 ETMERRENRLKDDIOTKSQIOQADKILELEKEKREACVSAOHLVHLKQEQHYBEKI 751
DB 249 ETMERRENRLKDDIOTKSEQIOQADKILELEKEKREACVSAOHLVHLKQEQHYBEKI 308
QY 752 KVLNQIKKDLADKETTLENMQRHEEFAHEKGIKILSEQKAMINAMDSKIRSLQRIVELS 811
DB 309 KVLNQIKKDLADKESLENMQRHEEFAHEKGIKILSEQKAMINAMDSKIRSLQRIVELS 368
QY 812 EANKLAANSSLETORNKACQEMISELROCKFYLETOAGKLEAQRNKLEEQLEKISHODH 871
DB 369 EANKLAANSSLETORNKACQEMISELROCKFYLETOAGKLEAQRNKLEEQLEKISHODH 428
QY 872 SDKNRLELETRLREVSLHEHEKQLEKQLTELQLSQERESQLTALQARAALESQRL 931
DB 429 SDKSRLELETRLREVSLHEHEKQLEKQLTELQLSQERESQLTALQARAALESQRL 488
QY 932 QAKTELETTAAEBBIOALTARDEIQKFPALRNSCTVIITDLEBQNLQNTEDNAELNN 991
DB 489 QAKTELETTAAEBBIOALTARDEIQKFPALRNSCTVIITDLEBQNLQNTEDNAELNN 548
QY 992 QNFYLSKQDLDEASGANDEIVQLRSEVDHLRRRITITERMOLTSQKOTMEALKTTTCM1BEQ 1051
DB 549 QNFYLSKQDLDEASGANDEIVQLRSEVDHLRRRITITERMOLTSQKOTMEALKTTTCM1BEQ 608

QY 1052 VMDLEALNDELLEKERQWEAWRSVLGDEKQSFECRVRELQRLMDTEKQSRARADORITES 1111
DB 609 VLDLEALNDELLEKERQWEAWRSVLGDEKQSFECRVRELQRLMDTEKQSRARADORITES 668
QY 1112 ROVVELAVKHEKAEIILALQOALKEOKLKAESLSDKLNDEKKHAMLEMNARSLOOKLETE 1171
DB 669 ROVVELAVKHEKAEIILALQOALKEOKLKAESLSDKLNDEKKHAMLEMNARSLOOKLETE 728
QY 1172 RELKORLLEBQAKLQQQMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYQLENIQVLY 1231
DB 729 RELKORLLEBQAKLQQQMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYQLENIQVLY 788
QY 1232 SHEKVMEGTIISOOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVFLQYNELKAL 1291
DB 789 SHEKVMEGTIISOOTKLIDFLQAKMDQPAKKK-----VFLQYNELKAL 833
QY 1292 EKEKARCAELEALOKTRIELRSAREAAHRKATDHPHPSTPATARQIAMSIVRSPEH 1351
DB 834 EKEKARCAELEALOKTRIELRSAREAAHRKATDHPHPSTPATARQIAMSIVRSPEH 893
QY 1352 QPSAMSLAPPPSSRRKESSTPEEFSRRLKERMHNIPHRNVGLNMRATKCAVCLDTVHF 1411
DB 894 QPSAMSLAPPPSSRRKESSTPEEFSRRLKERMHNIPHRNVGLNMRATKCAVCLDTVHF 953
QY 1412 GRQASKCLEQVMCHPKCSTCLPATCGLPAEYATHETEAFCRDKNMSPGLQTEPSSSLH 1471
DB 954 GRQASKCLEQVMCHPKCSTCLPATCGLPAEYATHETEAFCRDKNMSPGLQTEPSSSLH 1013
QY 1472 LEGMMKVPNNKRGQCGWDRKVIIVLEGSKVLIIYDNEAREAGQRPVEFECLPDGDVSIH 1531
DB 1014 LEGMMKVPNNKRGQCGWDRKVIIVLEGSKVLIIYDNEAREAGQRPVEFECLPDGDVSIH 1073
QY 1532 GAVGASLANATA----- 1545
DB 1074 GAVGASLANATAKADVPYIILKMESHPTTCWPGRTLYLLAPSPDKQKRWYTALESVVAGG 1133
QY 1546 ----EKAEDAALLGNLSLLLEGDDRLDMNCTLPFSDQVVLVGTGEGLYALNVLKNSLTH 1601
DB 1134 RVSEKAEADAALLGNLSLLLEGDDRLDMNCTLPFSDQVVLVGTGEGLYALNVLKNSLTH 1193
QY 1602 VPGIGAVFYIYIKOLEKLLMIAGEERALCLVDVKVKQSLAOSHJPAQDPDISNIFEAV 1661
DB 1194 IPGIGAVFYIYIKOLEKLLMIAGEERALCLVDVKVKQSLAOSHJPAQDPDVSPNIFEAV 1253
QY 1662 KGCHLFAGKIENTGLCIICAAMPKSVILRYNENLSKYCIKREIETSEPCSCIHFTNYSIL 1721
DB 1254 KGCHLFAGKIENTGLCIICAAMPKSVILRYNENLSKYCIKREIETSEPCSCIHFTNYSIL 1313
QY 1722 IGTNKFYEIDMKQYTLBEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREBYLLCFH 1781
DB 1314 IGTNKFYEIDMKQYTLBEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREBYLLCFH 1373
QY 1782 EFGVFDVSYGRSRSTDLLKWSRLPLAPAYREPVLVTHFNLSLEVIEIQARSSAGTPARAY 1841
DB 1374 EFGVFDVSYGRSRSTDLLKWSRLPLAPAYREPVLVTHFNLSLEVIEIQARSSAGTPARAY 1433
QY 1842 LDIPNRYLGPALSSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPTSRSSPNKRG 1901
DB 1434 LEIPNRYLGPALSSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPTSRSSPNKRG 1493
QY 1902 PPTVNEHITKRVASSPAPPEGSPHREPSTPHRY--REGTELRDRKSPGRPLEREKSPG 1959
DB 1494 PPTVNEHITKRVASSPAPPEGSPHREPSTPHRYDRREGTELRDRKSPGRPLEREKSPG 1553
QY 1960 RMLSTRERSGRLPEDSSRGRLPAGAVRTPLSQVNNKRGQSA 2002
DB 1554 RMLSTRERSGRLPEDSSRGRLPAGAVRTPLSQVNNKWDQSS 1596

RESULT 14
US-10-017-216-6
; Sequence 6, Application US/10017216
; Publication No. US20020160483A1

; GENERAL INFORMATION:  
 ; APPLICANT: KAPPELLER-LIBERMANN, ROSANA  
 ; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prot  
 ; TITLE OF INVENTION: Kinase and Uses Therefor  
 ; FILE REFERENCE: 10147-57U1  
 ; CURRENT APPLICATION NUMBER: US/10/017,216  
 ; PRIOR FILING DATE: 2001-10-23  
 ; PRIOR APPLICATION NUMBER: US 60/242,429  
 ; PRIOR FILING DATE: 2000-10-23  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 1597  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-017-216-6

Query Match 71.4%; Score 7491.5; DB 13; Length 1597;  
 Best Local Similarity 93.1%; Pred. No. 0;  
 Matches 1492; Conservative 16; Mismatches 12; Indels 83; Gaps 4;

Qy	468	MEQEMTLRLRRVSEVAVLSQKEVELKASQETQRSLLLEQDLATVITCSSLKRSLEQARME	527
Db	9	MEQEMTLRLRRVSEVAVLSQKEVELKASQETQRSLLLEQDLATVITCSSLKRSLEQARME	68
Qy	528	VSQEDDKALQLLHDIREQSRKLEIKEQFYQAOVEEMRLMMNQLBEDLVSARRSDLYES	587
Db	69	VSQEDDKALQLLHDIREQSRKLEIKEQFYQAOVEEMRLMMNQLBEDLVSARRSDLYES	128
Qy	588	ELRESLAAEFKRXKATECOHLLKAKDOQKPVGYAKLEKINABQQLKIQBLQEKLEK	647
Db	129	ELRESLAAEFKRXKANECHOLMKAKDOQKPVGYKLEKINABQQLKIQBLQEKLEK	188
Qy	648	A-----AKERARELEKIQNRDSSSEGIKKLVABERRHLENKVKRL	691
Db	189	AVKASTEATELLQNRQKERAERELEKLNHRDSSSEGIKKLVABERRHLENKVKRL	248
Qy	692	ETWERENRLKDDIQTKSQOIQOMADKIILEBEKHREAQVSAQHLEVHLKQEQHYEEKI	751
Db	249	ETWERENRLKDDIQTKSQOIQOMADKIILEBEKHREAQVSAQHLEVHLKQEQHYEEKI	308
Qy	752	KVLDNQIKKOLADKETLENMMQREBEAHEKGKILSEQKAMINAMDSKIRSLRQIVELS	811
Db	309	KVLDNQIKKOLADKESLENMMQREBEAHEKGKILSEQKAMINAMDSKIRSLRQIVELS	368
Qy	812	EANKLAANSLSLFRNNKQAEEMISLRQOKFVLETQAGKLEAQNKLBEQLEKISHQDH	871
Db	369	EANKLAANSLSLFRNNKQAEEMISLRQOKFVLETQAGKLEAQNKLBEQLEKISHQDH	428
Qy	872	SDKNRLLLETRLRVLSLEHEQKLEKQLTELQLSLOERESQLTALQAAARAALESQLR	931
Db	429	SDKSRLLLETRLRVLSLEHEQKLEKQLTELQLSLOERESQLTALQAAARAALESQLR	488
Qy	932	QAKTELEETAEABESIQALTARDBIQKFDALRNSCTVITDLBEQLNQLTQEDNAELNN	991
Db	489	QAKTELEETAEABESIQALTARDBIQKFDALRNSCTVITDLBEQLNQLTQEDNAELNN	548
Qy	992	QNFYLSKQLDEASGANDETVQLRSEVDHLRREITEREMOLTSKOTMEALKTKTCTMLEEQ	1051
Db	549	QNFYLSKQLDEASGANDETVQLRSEVDHLRREITEREMOLTSKOTMEALKTKTCTMLEEQ	608
Qy	1052	VMDLEALNDELLEKPERQAWRSVLGDEKSEQFECRVRELQRLMDTEKOSRAPADQRITES	1111
Db	609	VLDLEALNDELLEKPERQAWRSVLGDEKSEQFECRVRELQRLMDTEKOSRAPADQRITES	668
Qy	1112	RQVVELAVKEHKAIEIALQALKEQKLEKASLSDDKLNDELKHAMLENNARSLQOKLETE	1171
Db	669	RQVVELAVKEHKAIEIALQALKEQKLEKASLSDDKLNDELKHAMLENNARSLQOKLETE	728
Qy	1172	RELKORLLEBEQAKLQOQMDLQKNHIFRLTQGLQEALDRADLLKTERSOLLEYQLENIQVLY	1231
Db	729	RELKORLLEBEQAKLQOQMDLQKNHIFRLTQGLQEALDRADLLKTERSOLLEYQLENIQVLY	788

Qy	1232	SHEKVKMEGTISQOTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALPTQVFLQYNELKLAL	1291
Db	789	SHEKVKMEGTISQOTKLIDFLQAKMDQPAKKKK-----VPLQYNELKLAL	833
Qy	1292	EKEKARCAELBEALOKTRIILRSAREEAHRKATDHPHSTPATARQOIAMSALVSPSH	1351
Db	834	EKEKARCAELBEALOKTRIILRSAREEAHRKATDHPHSTPATARQOIAMSALVSPSH	893
Qy	1352	QPSAMSLIAPPSSRRKESSTPEEFSSRLKERMHNIIPHRFNVGLNMRATKCAVCLDTVHF	1411
Db	894	QPSAMSLIAPPSSRRKESSTPEEFSSRLKERMHNIIPHRFNVGLNMRATKCAVCLDTVHF	953
Qy	1412	GRQASKLCLEQVMCHPKCSTCLPATCGLPAEYATHFTFAFCRDKMNSPGLOTKEPSSSLH	1471
Db	954	GRQASKLCLEQVMCHPKCSTCLPATCGLPAEYATHFTFAFCRDKMNSPGLOTKEPSSSLH	1013
Qy	1472	LEGMMKVPNNKRGQOQWDRKYIVLEGSKVLIIYDNEAREAGQRPVEFEFELCLPDGDVSIH	1531
Db	1014	LEGMMKVPNNKRGQOQWDRKYIVLEGSKVLIIYDNEAREAGQRPVEFEFELCLPDGDVSIH	1073
Qy	1532	GAVGASELANATA-----	1545
Db	1074	GAVGASELANATAKADVPYILKMESHPHHTTCWPGRTLYLLAPSPDQQRWVTALESVVAGG	1133
Qy	1546	-----EKAEADAALLGNLSLLKLEGGDRLDNMCTLPFSDQVVLVGTTEGLYALNVLKNSLTH	1601
Db	1134	RVSREKAADAALLGNLSLLKLEGGDRLDNMCTLPFSDQVVLVGTTEGLYALNVLKNSLTH	1193
Qy	1602	VPGIGAVFQYIIKDLKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDIPSNIFEAV	1661
Db	1194	IPGIGAVFQYIIKDLKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDIPSNIFEAV	1253
Qy	1662	KGCHLFGAGKIENGLCI CAAMPKSVILRYNENLSKYCIKEIETSEPCSHFTNYSIL	1721
Db	1254	KGCHLFGAGKIENGLCI CAAMPKSVILRYNENLSKYCIKEIETSEPCSHFTNYSIL	1313
Qy	1722	IGTNKFEIDMKQYTLLEFLLDKNDHSLAPAVFAASSNSFPVSVQVNSAQOREEYLLCFH	1781
Db	1314	IGTNKFEIDMKQYTLLEFLLDKNDHSLAPAVFAASSNSFPVSVQVNSAQOREEYLLCFH	1373
Qy	1782	EFQVFDVSYGRSRDLDLKWRLPLAFAYREPILFVTHFNSLEVIEIQARSSAGTPARAY	1841
Db	1374	EFQVFDVSYGRSRDLDLKWRLPLAFAYREPILFVTHFNSLEVIEIQARSSAGTPARAY	1433
Qy	1842	LDIENPRYLGPATSSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPSTSSSPNKG	1901
Db	1434	LEIENPRYLGPATSSGAIYLAASSYQDKLRVICCKGNLVKESGTEHHRGPSTSSSPNKG	1493
Qy	1902	PPTYNEHITKRVASSPAPPEGSHPREPSTPHRY--REGTELRDRKSPGRPEREKSPG	1959
Db	1494	PPTYNEHITKRVASSPAPPEGSHPREPSTPHRYDRREGTELRDRKSPGRPEREKSPG	1553
Qy	1960	RMUSTRERSPGLFEDSSRGLPAGAVRTPLSQVNNKRGOSA 2002	
Db	1554	RMUSTRERSPGLFEDSSRGLPAGAVRTPLSQVNNKRGOSA 1596	

RESULT 15  
 US-10-412-897-3  
 ; Sequence 3, Application US/10412897  
 ; Publication No. US2003020224A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bristol-Myers Squibb Company  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES ENCODING THE HUMAN CITRON KINASE  
 ; TITLE OF INVENTION: POLYPEPTIDE, BMSNKC\_0020/0021  
 ; FILE REFERENCE: D0193 NP  
 ; CURRENT APPLICATION NUMBER: US/10/412,897  
 ; PRIOR FILING DATE: 2003-04-11  
 ; PRIOR APPLICATION NUMBER: U.S. 60/372,745  
 ; NUMBER OF SEQ ID NOS: 77  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 3

QY 1518 EFELCLPDGDSVHGAVGASELANTAKA-----1545  
Db 904 EFELCLPDGDSVHGAVGASELANTAKADVPYILKMESHPTTCWPGRTYLLAPSPFDK 963  
QY 1546 -----EKAEDAKLLGNSLLKLEGGDDRLDMNCTLPFSDQVVLVGTTEE 1587  
Db 964 QRWVTALESVVAGGRVSRKAEADAKLLGNSLLKLEGGDDRLDMNCTLPFSDQVVLVGTTEE 1023  
QY 1588 GLYALNVKNSLTHVPGIGAVFOYIIKDI-EKLLMTAGEERALCLVDVKKVQSLAOSHL 1647  
Db 1024 GLYALNVKNSLTHVPGIGAVFOYIIKDI-EKLLMTAGEERALCLVDVKKVQSLAOSHL 1083  
QY 1648 PAQPDISPNI FEAVKGCCHLFCAGKIENGLCICCAAMPKSVVILRYNENLSKYCIKKEIETS 1707  
Db 1084 PAQPDISPNI FEAVKGCCHLFCAGKIENGLCICCAAMPKSVVILRYNENLSKYCIKKEIETS 1143  
QY 1708 BPCSHFTNYSILIGTNKFEYIDMKOYTLEBFLDKNDHSLAPFAASSNSPVSIVQV 1767  
Db 1144 EPCSHFTNYSILIGTNKFEYIDMKOYTLEBFLDKNDHSLAPFAASSNSPVSIVQV 1203  
QY 1768 NSAQOREYLLCPHEFGVFDVSYGRRSRTDLDLKWRLPLAFAYREPFLFVTHFNSLEVE 1827  
Db 1204 NSAQOREYLLCPHEFGVFDVSYGRRSRTDLDLKWRLPLAFAYREPFLFVTHFNSLEVE 1263  
QY 1828 IQARSSAGTPARAYLDIPNPRYLGPALISSGAIYLAASYQDKLAVICCKGNLVKESGTEHH 1887  
Db 1264 IQARSSAGTPARAYLDIPNPRYLGPALISSGAIYLAASYQDKLAVICCKGNLVKESGTEHH 1323  
QY 1888 RGPSTSSSPNKGPPTYNESHITKRVASSPAPPEGSPHREPSTPHRY--REGTELRD 1945  
Db 1324 RVPSTSSSPNKGPPTYNESHITKRVASSPAPPEGSPHREPSTPHRYRDRGRTELRD 1383  
QY 1946 KSPGRPLEREKSPGRMLSTRERSPGSLFEDSSRGRLPAGAVRTPLSQVNVKRGQSA 2002  
Db 1384 KSPGRPLEREKSPGRMLSTRERSPGSLFEDSSRGRLPAGAVRTPLSQVNVKRWQSS 1440

Search completed: July 3, 2004, 10:26:03  
Job time : 110 secs

QY 574 DLVSARRRSDLYESLRESRLAABEFKPKATCQHKLLKAKDQGPVEGVAKLEKINAE 633  
Db 1 DLVSARRRSDLYESLRESRLAABEFKPKATCQHKLLKAKDQGPVEGVAKLEKINAE 60  
QY 634 QOLKIQOELQEKLEKA-----AKERAERELEKLNKREDSSEGIKKIYEA 677  
Db 61 QOLKIQOELQEKLEKAVKASTEATELLQNIQAKERAERELEKLNKREDSSEGIKKIYEA 120  
QY 678 BERRSHLENKVRLETMERENRLKDDIQTKSQIQQWADKILELEKHXREAQVSAQHLE 737  
Db 121 E-----ELEEXHREAQVSAQHLE 138  
QY 738 VHLKQKQHYEBKIKVLDNQIKKDLADKETLNMQRHEEBAHEKGIKILSEQKAMINAMD 797  
Db 139 VHLKQKQHYEBKIKVLDNQIKKDLADKESLNMQRHEEBAHEKGIKILSEQKAMINAMD 198  
QY 798 SKIRSLQRIVELSEBANKLAANSSLTQRNMKAQEEI-SELQKQFYLETQAGKLEAQN 857  
Db 199 SKIRSLQRIVELSEBANKLAANSSLTQRNMKAQEEI-SELQKQFYLETQAGKLEAQN 258  
QY 858 KLEEQLEKISHQCHSDKNRLLLEFRLRVSLEHEEQKLEKRLQTELOLSQRESOLT 917  
Db 259 KLEEQLEKISHQCHSDKNSLLEFRLRVSLEHEEQKLEKRLQTELOLSQRESOLT 318  
QY 918 ALQARAALLESQLOAKTELETTTAAEIEIQALTAHRDEIQKFDALRNSCTVITDLEE 977  
Db 319 ALQARAALLESQLOAKTELETTTAAEIEIQALTAHRDEIQKFDALRNSCTVITDLEE 378  
QY 978 QLNQLTEDNAELNNQNFYLSKQLDASGANDIIVQIRSEVDHLRREITERENMOLTSQQT 1037  
Db 379 QLNQLTEDNAELNNQNFYLSKQLDASGANDIIVQIRSEVDHLRREITERENMOLTSQQT 438  
QY 1038 MEALKTCTMLEEQVNDLEALDELLEKERQWAEWSVLGDEKSFQECRVRELQRLMDTE 1097  
Db 439 MEALKTCTMLEEQVNDLEALDELLEKERQWAEWSVLGDEKSFQECRVRELQRLMDTE 498  
QY 1098 KQSRARADQRI TESQVVELAVKEHKAETILALQALKEQKLAESLSDKINDLEKXHML 1157  
Db 499 KQSRARADQRI TESQVVELAVKEHKAETILALQALKEQKLAESLSDKINDLEKXHML 558  
QY 1158 ENNARSLOQKLETERELKQRLLEQAKLOQOQMDLQKNHIFRLTQGLQALDRADLLKTER 1217  
Db 559 ENNARSLOQKLETERELKQRLLEQAKLOQOQMDLQKNHIFRLTQGLQALDRADLLKTER 618  
QY 1218 SDLEYQLENIQVLYSHEKVMGEGTISQQTKLIDFLQAKMDQPAKKKGIFGRRKEDPALP 1277  
Db 619 SDLEYQLENIQVLYSHEKVMGEGTISQQTKLIDFLQAKMDQPAKKK-----665  
QY 1278 TOVPIQYNELKUALKEKARCAELBEALQKTRIELRSAREEAAHRAKATDHPHPSTPATAR 1337  
Db 666 --VPLQYNELKUALKEKARCAELBEALQKTRIELRSAREEAAHRAKATDHPHPSTPATAR 723  
QY 1338 QOIAMSIVRSPEHOPFSAMSLIAPPSSRREKESSTPEEFSSRLKERMHNI PHRFNVGLNM 1397  
Db 724 QOIAMSIVRSPEHOPFSAMSLIAPPSSRREKESSTPEEFSSRLKERMHNI PHRFNVGLNM 783  
QY 1398 RATKCAVCLDVTVHFGQKQKLECEQVMCHPKCSTCLPATCGLPAEYATHFTFAFCRDXKN 1457  
Db 784 RATKCAVCLDVTVHFGQKQKLECEQVMCHPKCSTCLPATCGLPAEYATHFTFAFCRDXKN 843  
QY 1458 SPGLQTKEPSSIHLEGMKVPNNKRGQOQNDKRYIVLEGSKVLIYDNEAREAQRPVE 1517  
Db 844 SPGLQTKEPSSIHLEGMKVPNNKRGQOQNDKRYIVLEGSKVLIYDNEAREAQRPVE 903

This Page Blank (uspto)

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2004, 09:41:14 ; Search time 2580 Seconds  
(without alignments)  
12292.380 Million cell updates/sec

Title: US-10-017-216-1

Perfect score: 6574

Sequence: 1 agagccgacagtgaggagat.....atcgagaatgaggtttaga 6574

Scoring table: IDENTIFY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6574	100.0	6574	14	US-10-017-216-1
2	6574	100.0	6574	15	US-10-325-430-10
3	6162	93.7	6162	15	US-10-325-430-11
4	6159	93.7	6159	14	US-10-017-216-3
5	5684.4	86.5	6298	13	US-10-415-011-43
6	5661.4	86.1	6165	14	US-10-028-946-1
7	5631	85.7	6189	13	US-10-028-946-1
8	5631	85.7	6189	13	US-10-262-511-1
9	5629.8	85.6	6201	13	US-09-964-956-10
10	5629.8	85.6	6201	13	US-10-262-511-13
11	5358	81.5	5877	14	US-09-964-956-8
12	2564.8	39.0	3131	13	US-10-028-946-3
13	2414	36.7	2542	13	US-10-276-774-137
14	2317.2	35.2	2497	13	US-10-262-511-7

15	1841.2	28.0	1870	13	US-10-262-511-3	Sequence 3, Appli
16	1524.2	23.2	1870	13	US-10-262-511-9	Sequence 9, Appli
17	1467.6	22.3	1915	13	US-10-262-511-11	Sequence 11, Appli
18	1406.4	21.4	2066	17	US-10-311-034-33	Sequence 33, Appli
19	1397.8	21.3	1515	9	US-09-804-471A-1	Sequence 1, Appli
20	1397.8	21.3	1515	15	US-10-238-709-1	Sequence 1, Appli
21	1397.8	21.3	1515	17	US-10-724-594-1	Sequence 1, Appli
22	965.6	14.7	2162	16	US-10-120-988-419	Sequence 419, App
23	950.6	14.5	957	16	US-10-412-897-1	Sequence 1, Appli
24	768	11.7	1133	15	US-10-282-048-1	Sequence 1, Appli
25	701.4	10.7	995	13	US-10-276-774-255	Sequence 255, App
26	421.4	6.4	2870	13	US-10-425-114-26241	Sequence 26241, A
27	350.4	5.3	354	14	US-10-040-739-9	Sequence 9, Appli
28	308	4.7	308	9	US-09-998-598-2328	Sequence 2328, Ap
29	306.4	4.7	308	9	US-09-815-343-132	Sequence 132, App
30	306.4	4.7	308	9	US-09-998-598-1929	Sequence 1929, Ap
31	306.4	4.7	308	13	US-10-097-105-132	Sequence 132, App
C 32	296	4.5	297	9	US-09-796-832-3885	Sequence 3885, Ap
C 33	296	4.5	297	15	US-10-040-862-3885	Sequence 3885, Ap
C 34	296	4.5	297	16	US-10-057-475B-3885	Sequence 3885, Ap
C 35	296	4.5	297	16	US-10-154-884B-3885	Sequence 3885, Ap
C 36	258	3.9	258	16	US-10-305-720-513	Sequence 513, App
37	243.6	3.7	1530	13	US-10-415-011-24	Sequence 24, Appli
38	243.6	3.7	4698	17	US-10-702-496-1	Sequence 1, Appli
39	243.6	3.7	5373	13	US-10-333-314-40	Sequence 40, Appli
40	236.8	3.6	6335	16	US-10-388-934-85	Sequence 85, Appli
41	234	3.6	5373	17	US-10-702-496-5	Sequence 5, Appli
42	228.2	3.5	2785	13	US-10-342-887-930	Sequence 930, App
43	228.2	3.5	2785	13	US-10-172-118-930	Sequence 930, App
44	228.2	3.5	5694	13	US-10-362-892-26	Sequence 26, Appli
45	228.2	3.5	5694	16	US-10-288-798-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1

US-10-017-216-1  
; Sequence 1, Application US/10017216  
; Publication No. US20020160483A1  
; GENERAL INFORMATION:  
; APPLICANT: KAPELIER-LIBERMANN, Rosana  
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prote  
; TITLE OF INVENTION: Kinase and Uses Therefor  
; FILE REFERENCE: 10147-57U1  
; CURRENT APPLICATION NUMBER: US/10/017,216  
; PRIOR FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: US 60/242,429  
; PRIOR FILING DATE: 2000-10-23  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 6574  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-216-1

Query Match	100.0%	Score 6574;	DB 14;	Length 6574;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 6574;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGAGCCGACAGTGGGAGATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCT	60	
Db	1	AGAGCCGACAGTGGGAGATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCT	60	
Qy	61	GGTGCTGTGAACCCATTTGCCAGCCGGGCTCCAGCTGAATCTGTTCCTCCAGGGGAAA	120	
Db	61	GGTGCTGTGAACCCATTTGCCAGCCGGGCTCCAGCTGAATCTGTTCCTCCAGGGGAAA	120	
Qy	121	CCACCTTTATGACTCAACAGCAGATGTCTCTCTTTCCGGAAGGATATTAGATGCC	180	
Db	121	CCACCTTTATGACTCAACAGCAGATGTCTCTCTTTCCGGAAGGATATTAGATGCC	180	

181	CTCTTTGTTCTCTTTTGAAGATGCACTCAGCTGCTCTGATGAAGATTTAAGCACGTTGAGC	240
181	CTCTTTGTTCTCTTTTGAAGATGCACTCAGCTGCTCTGATGAAGATTTAAGCACGTTGAGC	240
241	AACTTTGTCGGGAAGTATTTCCGACACATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCA	300
241	AACTTTGTCGGGAAGTATTTCCGACACATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCA	300
301	AAAGCACTTCCAAAGTCAGAAAGTCTTTGAGTTTGTTGTTCACTTTGCTGAAGTCAGGTGGTA	360
301	AAAGCACTTCCAAAGTCAGAAAGTCTTTGAGTTTGTTGTTCACTTTGCTGAAGTCAGGTGGTA	360
361	AGAGAGAAAGCACCGGGACATCTATGCTATGAAGAAGTATGAAGAAGAGGCTTTATTG	420
361	AGAGAGAAAGCAACCGGGACATCTATGCTATGAAGAAGTATGAAGAAGAGGCTTTATTG	420
421	GCCACGAGCAGGTTTTCATTTTTTTAGGAAGACGGGAACATATTTATCTCGAAGCACAAAGC	480
421	GCCACGAGCAGGTTTTCATTTTTTTAGGAAGACGGGAACATATTTATCTCGAAGCACAAAGC	480
481	CGGTGGAATCCCCAATTACAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAG	540
481	CGGTGGAATCCCCAATTACAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAG	540
541	GAATATCAGCTCGAGGGGACTTGCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGAT	600
541	GAATATCAGCTCGAGGGGACTTGCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGAT	600
601	GAATAACCTGATACAGTTTACCTPAGCTGAGCTGATTTTGGCTGTTTCAACGCGTTCACTG	660
601	GAATAACCTGATACAGTTTACCTPAGCTGAGCTGATTTTGGCTGTTTCAACGCGTTCACTG	660
661	ATCGGATAGCTGATCGAGACATCAAGCCTGAGCAATCTCGTTGACCCGACAGACAC	720
661	ATCGGATAGCTGATCGAGACATCAAGCCTGAGCAATCTCGTTGACCCGACAGACAC	720
721	ATCAAGCTGGTGGATTTTGGATCTGCGCGGAAAAATGAATTTCAAACAAGATGGTGAATGCC	780
721	ATCAAGCTGGTGGATTTTGGATCTGCGCGGAAAAATGAATTTCAAACAAGATGGTGAATGCC	780
781	AAACTCCGATTTGGACCCCAGATATCATGGCTCTTGAAGTCTGACTGTGATGAACGGG	840
781	AAACTCCGATTTGGACCCCAGATATCATGGCTCTTGAAGTCTGACTGTGATGAACGGG	840
841	GATGAAAAAGGCACCTACGGCCTGGACTGTGACTGGTGGTCACTGGGCGCTGATTGCCAT	900
841	GATGAAAAAGGCACCTACGGCCTGGACTGTGACTGGTGGTCACTGGGCGCTGATTGCCAT	900
901	GAGATGATTTATGGAGATCCCCCTTCGACAGAGGGAACCTCTCCGAGAACCTTCAATTAAC	960
901	GAGATGATTTATGGAGATCCCCCTTCGACAGAGGGAACCTCTCCGAGAACCTTCAATTAAC	960
961	ATTATGAATTTCCAGCGGTTTTTGAATTTTCCAGATGACCCCAAGATGACAGTGACTTT	1020
961	ATTATGAATTTCCAGCGGTTTTTGAATTTTCCAGATGACCCCAAGATGACAGTGACTTT	1020
1021	CTTGATCTGATTTCAAAGCTTTGTTGTGGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTT	1080
1021	CTTGATCTGATTTCAAAGCTTTGTTGTGGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTT	1080
1081	TGTTGCCATCTTTCTCTCTAAATTTGACTGGGAACAACATTCGTAACTCTCTCCCCCCC	1140
1081	TGTTGCCATCTTTCTCTCTAAATTTGACTGGGAACAACATTCGTAACTCTCTCCCCCCC	1140
1141	TTCGTTTCCCCACCTCAAGTCTGACGATGACACCTTCCAATTTTGATGAACACAGAGAAGAT	1200
1141	TTCTGTTCCCCACCTCAAGTCTGACGATGACACCTTCCAATTTTGATGAACACAGAGAAGAT	1200
1201	TCGTGGGTTTATCTCTCTCGTGGCAGCTGAGCCCTCAGGCTTCTCGGTGAGAACTG	1260
1201	TCGTGGGTTTATCTCTCTCGTGGCAGCTGAGCCCTCAGGCTTCTCGGTGAGAACTG	1260
1261	CCGTTTGTGGGGTTTTCGTACAGCAAGGCACCTGGGATTTCTGGTAGATCTGAGTCTGTT	1320

1261	CGTTTGTGGGTTTTTGGTACAGCAGGCATCGGGATCTTGGTAGATCTGAGTCTGTT	1320
1321	GTGTCGGGTCTGGAATCCCTCTGCCAAGCTAGCTCCATGGAAAAGAACTTCTCATCAAA	1380
1321	GTGTCGGGTCTGGAATCCCTCTGCCAAGCTAGCTCCATGGAAAAGAACTTCTCATCAAA	1380
1381	ACCAAGAGCTACAAGACTCTCAGGACCAAGTGTCACAAGATGGACGGAATGACCCGG	1440
1381	ACCAAGAGCTACAAGACTCTCAGGACCAAGTGTCACAAGATGGACGGAATGACCCGG	1440
1441	TTACATCTGGAGAGTGTACAGAGTGGAGGCTGTGCTTTAGTCAGAAGGAGTGGAGCTGAAG	1500
1441	TTACATCTGGAGAGTGTACAGAGTGGAGGCTGTGCTTTAGTCAGAAGGAGTGGAGCTGAAG	1500
1501	GGCTCTGAGACTCAGAGATCCCTCTGGACGAGGACCTTGCTACCTAGATCAGAGAATGC	1560
1501	GGCTCTGAGACTCAGAGATCCCTCTGGACGAGGACCTTGCTACCTAGATCAGAGAATGC	1560
1561	AGTAGCTTAAAGCGAAGTTTGGACGAAGCAGGATGGAGGTCTCCAGGAGGATGACAAA	1620
1561	AGTAGCTTAAAGCGAAGTTTGGACGAAGCAGGATGGAGGTCTCCAGGAGGATGACAAA	1620
1621	GCATCTGAGCTTCTCCATGATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAG	1680
1621	GCATCTGAGCTTCTCCATGATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAG	1680
1681	CAGGAGTACCAAGGCTCAAGTGGGAAGAAATGAGGTTGATCAATCACTTGGAAAGAGAT	1740
1681	CAGGAGTACCAAGGCTCAAGTGGGAAGAAATGAGGTTGATCAATCACTTGGAAAGAGAT	1740
1741	CTTGTCTCAGCAAGAAGCCGAGTGATCTCTACGAATCTGAGCTGAGAGATCTTCGGCTT	1800
1741	CTTGTCTCAGCAAGAAGCCGAGTGATCTCTACGAATCTGAGCTGAGAGATCTTCGGCTT	1800
1801	GCTGCTGAAGAAATTCAGCGGAAGCCAGAGATGTCAGCATTAATCTGTCGAAGGCTTAAG	1860
1801	GCTGCTGAAGAAATTCAGCGGAAGCCAGAGATGTCAGCATTAATCTGTCGAAGGCTTAAG	1860
1861	GATCAAGGGAAGCCTGAAGTGGGAGAAATATCGAAACTCGGAAGATCAATCTCTGAGCAG	1920
1861	GATCAAGGGAAGCCTGAAGTGGGAGAAATATCGAAACTCGGAAGATCAATCTCTGAGCAG	1920
1921	CAGCTCAAAATTCAGGAGCTCCAAAGAGAACTCGAGAAAGCTGCAAGAGGCGGAGCCGAG	1980
1921	CAGCTCAAAATTCAGGAGCTCCAAAGAGAACTCGAGAAAGCTGCAAGAGGCGGAGCCGAG	1980
1981	AGGGAGCTGGAGAGAGCTGCAGAACCGGAGAGATCTTCTTCTGAAGGATCAGAAAGAAGCTG	2040
1981	AGGGAGCTGGAGAGAGCTGCAGAACCGGAGAGATCTTCTTCTGAAGGATCAGAAAGAAGCTG	2040
2041	GTGGAAGCTGAGAAACGCGCGCATTTCTCTGAGAAACAAGTTAAAGAGCTAGAGACCATG	2100
2041	GTGGAAGCTGAGAAACGCGCGCATTTCTCTGAGAAACAAGTTAAAGAGCTAGAGACCATG	2100
2101	GAGCCTAGAGAAACAGACTGAAGATGACATCCAGACAAAATCCCAACAGATCCAGCAG	2160
2101	GAGCCTAGAGAAACAGACTGAAGATGACATCCAGACAAAATCCCAACAGATCCAGCAG	2160
2161	ATGGCTGATAAAATTTCTGAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTGACCCAG	2220
2161	ATGGCTGATAAAATTTCTGAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTGACCCAG	2220
2221	CACCTAGAAGTGCACCTGAAACAGAAAGACGACATATGAGGAAAGATTAAAGTTGTG	2280
2221	CACCTAGAAGTGCACCTGAAACAGAAAGACGACATATGAGGAAAGATTAAAGTTGTG	2280
2281	GACAAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAACATGATCAGAGA	2340
2281	GACAAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAACATGATCAGAGA	2340
2341	CAGGAGGAGGCCCATGAGAAGGGCAAAATTTCTCAGCGAACAGNACGATGATCAAT	2400



QY 4561 CCGGTGGAAGAAATTTGAGCTGTGCTTCCCGACGGGATGTATCTATTCAATGGTCCCGTT 4620  
Db 4561 CCGGTGGAAGAAATTTGAGCTGTGCTTCCCGACGGGATGTATCTATTCAATGGTCCCGTT 4620  
QY 4621 GGTGCTTCCGAATCTCGAATACAGCCAAAGCAGAAAAGCAGAAAGCTGATCTAAACTG 4680  
Db 4621 GGTGCTTCCGAATCTCGAATACAGCCAAAGCAGAAAAGCAGAAAGCTGATCTAAACTG 4680  
QY 4681 CTTTGGAAATCTCCCTGCTGAACTTGGAAAGGTGATGACCGTCTAGACATGAATGCACGCTG 4740  
Db 4681 CTTTGGAAATCTCCCTGCTGAACTTGGAAAGGTGATGACCGTCTAGACATGAATGCACGCTG 4740  
QY 4741 CCCTTCAGTGACAGAGTGTGTGTGGGACACGAGGAAGGGCTCTACGCCCTGAATGTC 4800  
Db 4741 CCCTTCAGTGACAGAGTGTGTGTGGGACACGAGGAAGGGCTCTACGCCCTGAATGTC 4800  
QY 4801 TTGAAAAATCTCCCTAACCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860  
Db 4801 TTGAAAAATCTCCCTAACCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860  
QY 4861 AAGGACCTGGAGAGTACTCATATGATAGCAGGAGAGCGGGCACTGTGTCTTGTGGAC 4920  
Db 4861 AAGGACCTGGAGAGTACTCATATGATAGCAGGAGAGCGGGCACTGTGTCTTGTGGAC 4920  
QY 4921 GTCAAGAAATGAAACAGTCCCTGGCCAGTCCACCTGCCCTGCCAGCCGACATCTCA 4980  
Db 4921 GTCAAGAAATGAAACAGTCCCTGGCCAGTCCACCTGCCCTGCCAGCCGACATCTCA 4980  
QY 4981 CCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTGGGCGAGGCAAGATTGAGAAC 5040  
Db 4981 CCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTGGGCGAGGCAAGATTGAGAAC 5040  
QY 5041 GGGCTCTGATCTGTGCGAGCCATGCCAGCAAGTGTCTTCTCGCTTACCAACGAAAC 5100  
Db 5041 GGGCTCTGATCTGTGCGAGCCATGCCAGCAAGTGTCTTCTCGCTTACCAACGAAAC 5100  
QY 5101 CTCAGCAAAATCTGCAATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5160  
Db 5101 CTCAGCAAAATCTGCAATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5160  
QY 5161 TTCAACAAATACAGTATCTCATTTGGAACCAATAAATTCTACGAATCGACATGAAGCAG 5220  
Db 5161 TTCAACAAATACAGTATCTCATTTGGAACCAATAAATTCTACGAATCGACATGAAGCAG 5220  
QY 5221 TACACGCTCGAGAAATCTGGATAGAAATGACATTCCTTGGCACTCTGCTGTGTGTC 5280  
Db 5221 TACACGCTCGAGAAATCTGGATAGAAATGACATTCCTTGGCACTCTGCTGTGTGTC 5280  
QY 5281 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGAGGTGAAACAGCGCAGGGCAGCGAGAG 5340  
Db 5281 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGAGGTGAAACAGCGCAGGGCAGCGAGAG 5340  
QY 5341 GAGTACTGCTGTGTTCACGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGAGCTAGC 5400  
Db 5341 GAGTACTGCTGTGTTCACGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGAGCTAGC 5400  
QY 5401 CGCACAGACGATCTCAAGTGGAGTCCCTTACCTTTTGGCCCTTTGCCTACAGAGAACCTTAT 5460  
Db 5401 CGCACAGACGATCTCAAGTGGAGTCCCTTACCTTTTGGCCCTTTGCCTACAGAGAACCTTAT 5460  
QY 5461 CTGTTTGTGACCCACTTCAACTCAGTCTGAAATTTGAGATTCAGGCAGGCTCCTCAGCA 5520  
Db 5461 CTGTTTGTGACCCACTTCAACTCAGTCTGAAATTTGAGATTCAGGCAGGCTCCTCAGCA 5520  
QY 5521 GGGACCCCTGCCGAGGCTACCTGGACATCCCGAACCCTGCTACCTGGGCCCTGCCAT 5580  
Db 5521 GGGACCCCTGCCGAGGCTACCTGGACATCCCGAACCCTGCTACCTGGGCCCTGCCAT 5580  
QY 5581 TCCTCAGGAGCGAATTTACTTGGGCTCTCATACGAGGATAAATTAAGGTCTATTGCTGC 5640  
Db 5581 TCCTCAGGAGCGAATTTACTTGGGCTCTCATACGAGGATAAATTAAGGTCTATTGCTGC 5640  
QY 5641 AAGGGAAACCTCGTGAAGGAGTCCGGCACTGAAACACCGGGGCCGCTCCACCTCCCGC 5700

Db 5641 AAGGGAAACCTCGTGAAGGAGTCCGGCACTGAAACACCGGGGCCGCTCCACCTCCGC 5700  
QY 5701 AGCAGCCCCAACAGAGGAGGCCACCCACCGTACAAAGAGCACAATCACCAGCGCTGGCC 5760  
Db 5701 AGCAGCCCCAACAGAGGAGGCCACCCACCGTACAAAGAGCACAATCACCAGCGCTGGCC 5760  
QY 5761 TCCAGCCACGCGCGCCGAGGCCACCGCAGCAGCCACCGCGAGAGCAACACCCACCGC 5820  
Db 5761 TCCAGCCACGCGCGCCGAGGCCACCGCAGCAGCCACCGCGAGAGCAACACCCACCGC 5820  
QY 5821 TACCGAGGGGGCGGACCGAGCTGCGCAGGCAAGTCTCTGGCCGCCCTTGGAGCGCA 5880  
Db 5821 TACCGAGGGGGCGGACCGAGCTGCGCAGGCAAGTCTCTCTGGCCGCCCTTGGAGCGCA 5880  
QY 5881 GAGAAGTCCCGCCCGGATGCTCAGCACGCGAGAGAGCGGTCCCGCGGAGGCTGTTT 5940  
Db 5881 GAGAAGTCCCGCCCGGATGCTCAGCACGCGAGAGAGCGGTCCCGCGGAGGCTGTTT 5940  
QY 5941 GAAGACAGCAGCGGGCCCGCTGCGGGAGCCGCTGAGGACCCCGCTGTCCACAGTG 6000  
Db 5941 GAAGACAGCAGCGGGCCCGCTGCGGGAGCCGCTGAGGACCCCGCTGTCCACAGTG 6000  
QY 6001 AACAAAGGAAAGAGGCGAGTGCCTCTCAAGTTTTCACGGTTTAACTGTCACTATTAT 6060  
Db 6001 AACAAAGGAAAGAGGCGAGTGCCTCTCAAGTTTTCACGGTTTAACTGTCACTATTAT 6060  
QY 6061 GACTGGAATAAAAGCTGGACACACCTGCCAGCTAACTGGTCACTCAGGATCATCCAG 6120  
Db 6061 GACTGGAATAAAAGCTGGACACACCTGCCAGCTAACTGGTCACTCAGGATCATCCAG 6120  
QY 6121 CTGAATCGAAGAAATCCCGGACAGGTTGAAAGTCTGTCTGAGAACAGATTATTCCTGA 6180  
Db 6121 CTGAATCGAAGAAATCCCGGACAGGTTGAAAGTCTGTCTGAGAACAGATTATTCCTGA 6180  
QY 6181 GCAGAGTTCATGTGACTTCTAGACGTGTGCTTAAAGTCTTAAAGTCTTAAAGTCTGAGC 6240  
Db 6181 GCAGAGTTCATGTGACTTCTAGACGTGTGCTTAAAGTCTTAAAGTCTTAAAGTCTGAGC 6240  
QY 6241 CAGCCACCTCTGCTTACAAAAGAGTACTTGTAGTCACTGACTGTAAAGAACTGTAA 6300  
Db 6241 CAGCCACCTCTGCTTACAAAAGAGTACTTGTAGTCACTGACTGTAAAGAACTGTAA 6300  
QY 6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCTGGAGCC 6360  
Db 6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCTGGAGCC 6360  
QY 6361 GAGAGCAATCTGTGTGTGTTTTGAGGACAGCAGCCACACTGTATTTAGTTCCA 6420  
Db 6361 GAGAGCAATCTGTGTGTGTTTTGAGGACAGCAGCCACACTGTATTTAGTTCCA 6420  
QY 6421 TAGCCAGGCTCAACAGGAGCAAGTGGCTGGCCCTTAAAAACACACAGATGACGAAATG 6480  
Db 6421 TAGCCAGGCTCAACAGGAGCAAGTGGCTGGCCCTTAAAAACACACAGATGACGAAATG 6480  
QY 6481 ATGTGTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGGATAGCAATTCATTT 6540  
Db 6481 ATGTGTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGGATAGCAATTCATTT 6540  
QY 6541 TTGGCTTTAAGAAAAATCGAAGATGTAGTTTGA 6574  
Db 6541 TTGGCTTTAAGAAAAATCGAAGATGTAGTTTGA 6574

## RESULT 2

US-10-325-430-10  
; Sequence 10, Application US/10325430  
; Publication No. US2003015325A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc  
; APPLICANT: Silos-Santiago, Inmaculada  
; APPLICANT: Rosenfeld, Julie Beth  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING

; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; TITLE OF INVENTION: 32838,336 and 52908
; FILE REFERENCE: MPI01-294PIRNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-325-430-10

		Query Match	100.0%;	Score 6574;	DB 15;	Length 6574;		
		Best Local Similarity	100.0%;	Pred. No. 0;				
		Matches 6574;	Conservative	0;	Mismatches	0;	Indels	0;
QY	1	AGAGCGCCAGTGGGAGATGTTGAAGTTCAAAATGAGAGCGCGGAATCCTTTGGATGCT	60					
DB	1	AGAGCGCCAGTGGGAGATGTTGAAGTTCAAAATGAGAGCGCGGAATCCTTTGGATGCT	60					
QY	61	GGTGCTGTGAACCATTCGCCAGCGGGCTCCAGSCTGAATCTCTTCCAGGGGAA	120					
DB	61	GGTGCTGTGAACCATTCGCCAGCGGGCTCCAGSCTGAATCTCTTCCAGGGGAA	120					
QY	121	CCACCCCTTATGACTCAACAGAGATGTCCTCTTTCCGGAAGGGATATTAGATGCC	180					
DB	121	CCACCCCTTATGACTCAACAGAGATGTCCTCTTTCCGGAAGGGATATTAGATGCC	180					
QY	181	CTCTTGTTCTCTTGAAGAAATGAGTACGCTGCTCTGTATGAAGATTAAGACGTGAGC	240					
DB	181	CTCTTGTTCTCTTGAAGAAATGAGTACGCTGCTCTGTATGAAGATTAAGACGTGAGC	240					
QY	241	AACCTTGTCCGGAAGTATTCGACACCATAGCTGAGTGTACAGAGCTCCAGCTTCGGCA	300					
DB	241	AACCTTGTCCGGAAGTATTCGACACCATAGCTGAGTGTACAGAGCTCCAGCTTCGGCA	300					
QY	301	AAGGACTTCGAAAGTCAGAAAGTCTTGTAGTGTGTCTACTTTGCTGAAAGTGGTGA	360					
DB	301	AAGGACTTCGAAAGTCAGAAAGTCTTGTAGTGTGTCTACTTTGCTGAAAGTGGTGA	360					
QY	361	ACAGAGAAAGCAACCGGGACATCTATGCTATGAAGTGTATGAAGAAAGGCTTTATTG	420					
DB	361	ACAGAGAAAGCAACCGGGACATCTATGCTATGAAGTGTATGAAGAAAGGCTTTATTG	420					
QY	421	GCCAGGAGCAGGTTTCATTTTTTGGGAAGAGCGGAACATATTATCTCGAAGCAACAGC	480					
DB	421	GCCAGGAGCAGGTTTCATTTTTTGGGAAGAGCGGAACATATTATCTCGAAGCAACAGC	480					
QY	481	CGTGATATCCCAATTACAGTATGCTTTTACAGACAAAATCACCTTTATCTGATGGAG	540					
DB	481	CGTGATATCCCAATTACAGTATGCTTTTACAGACAAAATCACCTTTATCTGATGGAG	540					
QY	541	GAATATCAGCTGGAGGGACTTGTCTACATTTTGAATAGATATGAGACCCAGTTAGAT	600					
DB	541	GAATATCAGCTGGAGGGACTTGTCTACATTTTGAATAGATATGAGACCCAGTTAGAT	600					
QY	601	GAAAACCTGATACAGTTTACCTAGCTAGCTGATTTTGGCTGTTCACAGCTTCATCTG	660					
DB	601	GAAAACCTGATACAGTTTACCTAGCTAGCTGATTTTGGCTGTTCACAGCTTCATCTG	660					
QY	661	ATGGGATACGTCATCGAGACATCAAGCTGAGACATTTCTGTTGACCGACAGACAC	720					
DB	661	ATGGGATACGTCATCGAGACATCAAGCTGAGACATTTCTGTTGACCGACAGACAC	720					
QY	721	ATCAAGCTGGTGGATTTGGATCTGCCGCGAAAATGAATTCAAAAGATGGTGAATGCC	780					
DB	721	ATCAAGCTGGTGGATTTGGATCTGCCGCGAAAATGAATTCAAAAGATGGTGAATGCC	780					
QY	781	AAACTCCCGATTTGGAGCCCGACATATACATGGCTCTGAGATGCTGATGATGACGG	840					

DB	781	AAACTCCCGATTTGGAGCCCGACAGATTACATGGCTCTCTGAAGTGTGACTGTGATGAACGG	840
QY	841	GATGAAAAGGCACCTACGCGCTGGACTGTGACTGTGTGTGCTAGTGGCGGTGATTGCCTAT	900
DB	841	GATGAAAAGGCACCTACGCGCTGGACTGTGACTGTGTGTGCTAGTGGCGGTGATTGCCTAT	900
QY	901	GAGATGATTTATGGGAGATCCCTCTCCAGAGGGAACCTCTGCGCAACCTTCAATAAC	960
DB	901	GAGATGATTTATGGGAGATCCCTCTCCAGAGGGAACCTCTGCGCAACCTTCAATAAC	960
QY	961	ATTATGAATTTCCAGCGGTTTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTT	1020
DB	961	ATTATGAATTTCCAGCGGTTTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTT	1020
QY	1021	CTTGATCTGATTTCAAAGCTTGTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTT	1080
DB	1021	CTTGATCTGATTTCAAAGCTTGTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTT	1080
QY	1081	TGCTGCCATCTTCT	1140
DB	1081	TGCTGCCATCTTCT	1140
QY	1141	TTCTTCCCACTTCAAGTCTGACGATGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1200
DB	1141	TTCTTCCCACTTCAAGTCTGACGATGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1200
QY	1201	TCGTGGGTTTCACT	1260
DB	1201	TCGTGGGTTTCACT	1260
QY	1261	CCGTCTTCTGGGTTTCTGATACAGCAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1320
DB	1261	CCGTCTTCTGGGTTTCTGATACAGCAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1320
QY	1321	GTGTCTGGGTTTCTGACT	1380
DB	1321	GTGTCTGGGTTTCTGACT	1380
QY	1381	AGCAAGAGCTACAGACT	1440
DB	1381	AGCAAGAGCTACAGACT	1440
QY	1441	TTATCTCGGAGTGTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAG	1500
DB	1441	TTATCTCGGAGTGTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAG	1500
QY	1501	GCCTCTGAGACTCAGAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1560
DB	1501	GCCTCTGAGACTCAGAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1560
QY	1561	AGTAGCTTAAAGCGAAGTTTGGAGCAAGCGGATGGAGGTGTTCCAGGAGGATGACAA	1620
DB	1561	AGTAGCTTAAAGCGAAGTTTGGAGCAAGCGGATGGAGGTGTTCCAGGAGGATGACAA	1620
QY	1621	GCACTGAGCTTCT	1680
DB	1621	GCACTGAGCTTCT	1680
QY	1681	CAGAGTACAGGCTCAAGTGGAGAAATCAGGTTGATGATGATCAGTTGGAGAGGAT	1740
DB	1681	CAGAGTACAGGCTCAAGTGGAGAAATCAGGTTGATGATGATCAGTTGGAGAGGAT	1740
QY	1741	CTTGCTCTCAGCAAGAACGAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1800
DB	1741	CTTGCTCTCAGCAAGAACGAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1800
QY	1801	CTTGCTCTCAGCAAGAACGAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1860
DB	1801	CTTGCTCTCAGCAAGAACGAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1860
QY	1861	GATCAAGGAGGCTGAGGAGGATATGCGAAATCTGAGAGGATCAATGCTGAGCAG	1920
DB	1861	GATCAAGGAGGCTGAGGAGGATATGCGAAATCTGAGAGGATCAATGCTGAGCAG	1920

1921 QY CAGCTCAAAATTCAGGAGCTCCAGAGAACTCGAGAAAGGCTCCAAAGAGCGAGCCGAG 1980  
1921 Db CAGCTCAAAATTCAGGAGCTCCAGAGAACTCGAGAAAGGCTCCAAAGAGCGAGCCGAG 1980  
1981 QY AGGAGCTGAGAAAGCTGCAAGACCGAGAGATTCCTTTGAAAGGCATCAGAAGAGAGCTG 2040  
1981 Db AGGAGCTGAGAAAGCTGCAAGACCGAGAGATTCCTTTGAAAGGCATCAGAAGAGAGCTG 2040  
2041 QY GTGGAAGCTGAGAAAGCCGCCATTCCTGCGAACAAGGTAAAGAGACTAGAGACCATG 2100  
2041 Db GTGGAAGCTGAGAAAGCCGCCATTCCTGCGAACAAGGTAAAGAGACTAGAGACCATG 2100  
2101 QY GAGCGTAGAAGAAACAGACTGAGAGATGACATCCAGACAAATCCCAACAGATCCAGAG 2160  
2101 Db GAGCGTAGAAGAAACAGACTGAGAGATGACATCCAGACAAATCCCAACAGATCCAGAG 2160  
2161 QY ATGGCTGATAAAATTCGAGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGCCAG 2220  
2161 Db ATGGCTGATAAAATTCGAGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGCCAG 2220  
2221 QY CACTAGAGTGCACCTGAACAGAAAGACGACACTATGAGAAAGATTAAGTGTG 2280  
2221 Db CACTAGAGTGCACCTGAACAGAAAGACGACACTATGAGAAAGATTAAGTGTG 2280  
2281 QY GACAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAAACATGATGAGAGA 2340  
2281 Db GACAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAAACATGATGAGAGA 2340  
2341 QY CAGCAGAGAGGCCCATGAGAAAGGCGAAATTCCTCAGCGAACAAGGCGATGATCAAT 2400  
2341 Db CAGCAGAGAGGCCCATGAGAAAGGCGAAATTCCTCAGCGAACAAGGCGATGATCAAT 2400  
2401 QY GCTATGATTCGAAGATCAGATCCCTGGAAACAGAGGATTTGGAACTGCTGAAGCCAT 2460  
2401 Db GCTATGATTCGAAGATCAGATCCCTGGAAACAGAGGATTTGGAACTGCTGAAGCCAT 2460  
2461 QY AAATCTGCAAAATAGCAGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATG 2520  
2461 Db AAATCTGCAAAATAGCAGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATG 2520  
2521 QY ATTCTGAACTCAGCAACAGAAATTTTACCTGGACACAGGCTGGAAAGTGGAGGCC 2580  
2521 Db ATTCTGAACTCAGCAACAGAAATTTTACCTGGACACAGGCTGGAAAGTGGAGGCC 2580  
2581 QY CAGAAACGAAAACTGGAGGAGAGCTGGAGAAAGATCAGCCACCAAGACACAGTGAACAAG 2640  
2581 Db CAGAAACGAAAACTGGAGGAGAGCTGGAGAAAGATCAGCCACCAAGACACAGTGAACAAG 2640  
2641 QY AATCGGCTGCTGAACTGAGACAGATTCGGGAGGTGAGTCTAGAGACAGAGAGCAG 2700  
2641 Db AATCGGCTGCTGAACTGAGACAGATTCGGGAGGTGAGTCTAGAGACAGAGAGCAG 2700  
2701 QY AAATCGAGCTCAAGCGCCAGCTCAGAGACTACAGCTCTCCCTGCGAGGCGCAGTCA 2760  
2701 Db AAATCGAGCTCAAGCGCCAGCTCAGAGACTACAGCTCTCCCTGCGAGGCGCAGTCA 2760  
2761 QY CAGTTGACAGCTCAGGCTGCAAGGCGGCCCTGGAGAGCAGCTTCGCCAGCGAAG 2820  
2761 Db CAGTTGACAGCTCAGGCTGCAAGGCGGCCCTGGAGAGCAGCTTCGCCAGCGAAG 2820  
2821 QY ACAGAGCTGGAAGAGACACAGCAGAAAGCTGAAGAGGAGATCCAGGCACTCAGCGCACAT 2880  
2821 Db ACAGAGCTGGAAGAGACACAGCAGAAAGCTGAAGAGGAGATCCAGGCACTCAGCGCACAT 2880  
2881 QY AGAGATGAAATCCAGCGCAAAATTTGATGCTCTTCTGTAACAGCTGATCTGTAATCAAGAC 2940  
2881 Db AGAGATGAAATCCAGCGCAAAATTTGATGCTCTTCTGTAACAGCTGATCTGTAATCAAGAC 2940  
2941 QY CTGGAGGAGCGCTAAACAGCTGACCGAGGACACGCTGAACTCAACAAACCAAACTTC 3000  
2941 Db CTGGAGGAGCGCTAAACAGCTGACCGAGGACACGCTGAACTCAACAAACCAAACTTC 3000

3001 QY TACTTGTCCAAACAACTCGATGAGGCTTCTGGCCCAACGACGAGATTGTACAACTCGCA 3060  
3001 Db TACTTGTCCAAACAACTCGATGAGGCTTCTGGCCCAACGACGAGATTGTACAACTCGCA 3060  
3061 QY AGTGAAGTGAACCATCTCCGCCGGAGATCAGGAAACGAGAGATGACGCTTACCAAGCCAG 3120  
3061 Db AGTGAAGTGAACCATCTCCGCCGGAGATCAGGAAACGAGAGATGACGCTTACCAAGCCAG 3120  
3121 QY AAGCAAAACGATGAGAGGCTCTGAAGACCACTGACCAATGCTGAGAGAACAGGTCATGGAT 3180  
3121 Db AAGCAAAACGATGAGAGGCTCTGAAGACCACTGACCAATGCTGAGAGAACAGGTCATGGAT 3180  
3181 QY TTGAGAGCCCTTAAACGATGAGCTGCTAGAAAAGAGCGGAGTGGAGGCTCGAGAGG 3240  
3181 Db TTGAGAGCCCTTAAACGATGAGCTGCTAGAAAAGAGCGGAGTGGAGGCTCGAGAGG 3240  
3241 QY GTCTGGGTGATGAGAAATCCAGTTTGTGTCGGGTTGAGAGCTGCGAGCTGCGAGAAATGCTG 3300  
3241 Db GTCTGGGTGATGAGAAATCCAGTTTGTGTCGGGTTGAGAGCTGCGAGCTGCGAGAAATGCTG 3300  
3301 QY GACACCGAGAAACAGAGCGGCGAGCGGATCAGCGGATCACCAGTCTCCCGAGGCTG 3360  
3301 Db GACACCGAGAAACAGAGCGGCGAGCGGATCAGCGGATCACCAGTCTCCCGAGGCTG 3360  
3361 QY GTGAGCTGCGCAGTGAAGGAGCAAGGCTGAGATTCTCGCTCTGCGAGAGGCTCTCAAA 3420  
3361 Db GTGAGCTGCGCAGTGAAGGAGCAAGGCTGAGATTCTCGCTCTGCGAGAGGCTCTCAAA 3420  
3421 QY GAGCAGAGCTGAAGCGCGAGAGGCTCTCTGACAAGCTCAATCACTGGAGAGAGCAT 3480  
3421 Db GAGCAGAGCTGAAGCGCGAGAGGCTCTCTGACAAGCTCAATCACTGGAGAGAGCAT 3480  
3481 QY GCTATGCTTGAATGAAATGCCAAGCTTACAGCAGAAAGCTGAGACTGAACGAGAGCTC 3540  
3481 Db GCTATGCTTGAATGAAATGCCAAGCTTACAGCAGAAAGCTGAGACTGAACGAGAGCTC 3540  
3541 QY AAAACAGAGGCTTCTGGAAGAGCAAGCCAAATTCAGCAGCAGATGGAACCTGCAAGAAAT 3600  
3541 Db AAAACAGAGGCTTCTGGAAGAGCAAGCCAAATTCAGCAGCAGATGGAACCTGCAAGAAAT 3600  
3601 QY CACATTTTCGCTGACTCAAGGACTGCAAGAGCTTAGATCGGCTGATCTACTGAAG 3660  
3601 Db CACATTTTCGCTGACTCAAGGACTGCAAGAGCTTAGATCGGCTGATCTACTGAAG 3660  
3661 QY ACAGAAAAGAGTGACTGGAGTATCAGCTGGAACCAATTCAGGTTCTCTATTCTCATGAA 3720  
3661 Db ACAGAAAAGAGTGACTGGAGTATCAGCTGGAACCAATTCAGGTTCTCTATTCTCATGAA 3720  
3721 QY AAGGTGAATGGAAGCCTATTTCTCAACAAACCAATCTCATTTTCTGCAAGCC 3780  
3721 Db AAGGTGAATGGAAGCCTATTTCTCAACAAACCAATCTCATTTTCTGCAAGCC 3780  
3781 QY AAAATGACCAACCTGCTAAAAAGAAAAGGGTTATTTAGTCGACGGAAAGAGGCCCT 3840  
3781 Db AAAATGACCAACCTGCTAAAAAGAAAAGGGTTATTTAGTCGACGGAAAGAGGCCCT 3840  
3841 QY GCTTTACCCACACAGGTTCTCTGCGAGTACAACTGAGCTGAGCTGGCCCTGGAGAGGAG 3900  
3841 Db GCTTTACCCACACAGGTTCTCTGCGAGTACAACTGAGCTGAGCTGGCCCTGGAGAGGAG 3900  
3901 QY AAAGCTCGCTGTGAGAGCTAGAGGAGCCCTTTCAGAAAGCCGCACTCAGAGCTCCGGTCC 3960  
3901 Db AAAGCTCGCTGTGAGAGCTAGAGGAGCCCTTTCAGAAAGCCGCACTCAGAGCTCCGGTCC 3960  
3961 QY GCCCGGAGGAGGCTGCCACCGCAAGCAACCGAACCAACCAACCAACCAACCAACCAAC 4020  
3961 Db GCCCGGAGGAGGCTGCCACCGCAACCGCAACCGAACCAACCAACCAACCAACCAAC 4020  
4021 QY ACCGCGAGGAGCAGAGATCGCCATGTCGCGCATGTCGCGTTCGCGAGAGCAGACCCAGT 4080  
4021 Db ACCGCGAGGAGCAGAGATCGCCATGTCGCGCATGTCGCGTTCGCGAGAGCAGACCCAGT 4080  
4081 QY GCCATGAGCTGTGCGCCCGCCCATCCAGCGCGCAGAAAGAGTCTTCACTCCAGAGGAA 4140

Db 4081 GGCATGAGCCTGCTGGCCCCGGCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGGAA 4140  
 Qy 4141 TTATGTCGGCGTCTTAAGGAAGCGATGACACCAATATTTCTTCAACCGATTTCAACGTAGGA 4200  
 Db 4141 TTATGTCGGCGTCTTAAGGAAGCGATGACACCAATATTTCTTCAACCGATTTCAACGTAGGA 4200  
 Qy 4201 CTGAACATCGGAGCACCAAAAGTGTGCTGTGTCTGAGATACCGTGCATTTGGAGCCGAG 4260  
 Db 4201 CTGAACATCGGAGCACCAAAAGTGTGCTGTGTCTGAGATACCGTGCATTTGGAGCCGAG 4260  
 Qy 4261 GCATCAAAATGTCGAAATGTCAAGTGTGTGTCAACCCCAAGTGTCCACGTGCTTGCCA 4320  
 Db 4261 GCATCAAAATGTCGAAATGTCAAGTGTGTGTCAACCCCAAGTGTCCACGTGCTTGCCA 4320  
 Qy 4321 GCCACTCGGGGTTGCTGCTGAAATGAGCCACACTTCAACCGAGGCTTTGCGGTGAC 4380  
 Db 4321 GCCACTCGGGGTTGCTGCTGAAATGAGCCACACTTCAACCGAGGCTTTGCGGTGAC 4380  
 Qy 4381 AAAATGAATCTCCAGAGTCTCCAGACCAAGGAGCCAGCAGCAGTTCACCTGGAAGG 4440  
 Db 4381 AAAATGAATCTCCAGAGTCTCCAGACCAAGGAGCCAGCAGCAGTTCACCTGGAAGG 4440  
 Qy 4441 TGGATGAAGTGCACGAGGATTAACAAAGAGGACAGCAAGGCTGGGACAGGAATACAT 4500  
 Db 4441 TGGATGAAGTGCACGAGGATTAACAAAGAGGACAGCAAGGCTGGGACAGGAATACAT 4500  
 Qy 4501 GTCCCTGGAGGATCAAAAGTCTCTATTTATGACATGAAGCCAGAGAGCTGGACAGAG 4560  
 Db 4501 GTCCCTGGAGGATCAAAAGTCTCTATTTATGACATGAAGCCAGAGAGCTGGACAGAG 4560  
 Qy 4561 CCGGTGGAAGAAATTTGAGTGTGCTTCCGACGGGGATGTATCTATTTATGCTGCTGCT 4620  
 Db 4561 CCGGTGGAAGAAATTTGAGTGTGCTTCCGACGGGGATGTATCTATTTATGCTGCTGCT 4620  
 Qy 4621 GGTGCTTCGAACTCGCAATACAGCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4680  
 Db 4621 GGTGCTTCGAACTCGCAATACAGCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4680  
 Qy 4681 CTTGGAATCTCCCTGCTGAACTGGAAGTGTATGACCTGCTAGACATGAATGCAAGCTG 4740  
 Db 4681 CTTGGAATCTCCCTGCTGAACTGGAAGTGTATGACCTGCTAGACATGAATGCAAGCTG 4740  
 Qy 4741 CCGCTTCAAGTACCAAGTGTGTGTGGGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4800  
 Db 4741 CCGCTTCAAGTACCAAGTGTGTGTGGGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4800  
 Qy 4801 TTGAAATCTCCCTGCTGAACTGGAAGTGTATGACCTGCTAGACATGAATGCAAGCTG 4860  
 Db 4801 TTGAAATCTCCCTGCTGAACTGGAAGTGTATGACCTGCTAGACATGAATGCAAGCTG 4860  
 Qy 4861 AAGGACCTGGAGAGCTACTATGATAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4920  
 Db 4861 AAGGACCTGGAGAGCTACTATGATAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4920  
 Qy 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4980  
 Db 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4980  
 Qy 4981 CCAACATTTTGAAGCTGCAAGGCTGCACTTTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5040  
 Db 4981 CCAACATTTTGAAGCTGCAAGGCTGCACTTTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5040  
 Qy 5041 GGGCTCTGATCTGTGAGCCATGCCCAGCAAGTGTGATTTCTCGGTACAAGCAAAAC 5100  
 Db 5041 GGGCTCTGATCTGTGAGCCATGCCCAGCAAGTGTGATTTCTCGGTACAAGCAAAAC 5100  
 Qy 5101 CTCAGCAATCTGATCCGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5160  
 Db 5101 CTCAGCAATCTGATCCGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5160  
 Qy 5161 TTCCAAATACAGTATCTTCAATGGAACCAATAAATTTCTAGAAATCCATGAAGCAG 5220

Db 5161 TTCCAAATACAGTATCTTCAATGGAACCAATAAATTTCTAGAAATCGAATCGAATCGAAGCAG 5280  
 Qy 5221 TACACGCTCGAGGAATTTCTGGATAAGAAATGACCAATTTCTTGGCACCTGCTGTGTTGCC 5280  
 Db 5221 TACACGCTCGAGGAATTTCTGGATAAGAAATGACCAATTTCTTGGCACCTGCTGTGTTGCC 5280  
 Qy 5281 GCCTTTCCCAACAGCTTTCCCTGTCTCAATCGTGCAGTGAACAGCCAGGAGGAGGAGGAGGAG 5340  
 Db 5281 GCCTTTCCCAACAGCTTTCCCTGTCTCAATCGTGCAGTGAACAGCCAGGAGGAGGAGGAGGAG 5340  
 Qy 5341 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGATGAGC 5400  
 Db 5341 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGATGAGC 5400  
 Qy 5401 CGCACAGACATCTCAAGTGGAGTGTGCTTACCTTTGGCCCTTTGGCTTACAGAGAACCTTAT 5460  
 Db 5401 CGCACAGACATCTCAAGTGGAGTGTGCTTACCTTTGGCCCTTTGGCTTACAGAGAACCTTAT 5460  
 Qy 5461 CTGTTTGTGACCACTTCAACTCACTCGAAGTAAATGAGATCCAGGACACCTCTCTAGCA 5520  
 Db 5461 CTGTTTGTGACCACTTCAACTCACTCGAAGTAAATGAGATCCAGGACACCTCTCTAGCA 5520  
 Qy 5521 GGGACCCCTGCCGAGCGTACCTGGACATCCCGAACCCCGCTACCTTGGCCCTTGGCCAT 5580  
 Db 5521 GGGACCCCTGCCGAGCGTACCTGGACATCCCGAACCCCGCTACCTTGGCCCTTGGCCAT 5580  
 Qy 5581 TCCTCAGAGCGATTTACTTTGGCGTCTCATACCGAGTAAATTAAGGGTCAATTTGCTGC 5640  
 Db 5581 TCCTCAGAGCGATTTACTTTGGCGTCTCATACCGAGTAAATTAAGGGTCAATTTGCTGC 5640  
 Qy 5641 AAGGAAACCTCTGTAAGAGGAGTCCGGACATGAACACACCCGGGGCCCGTCCACCTCCCGC 5700  
 Db 5641 AAGGAAACCTCTGTAAGAGGAGTCCGGACATGAACACACCCGGGGCCCGTCCACCTCCCGC 5700  
 Qy 5701 AGACGCCCCAACAGCGAGGCCACCCACCTACACGAGCACATCACCAGCGGTGGCC 5760  
 Db 5701 AGACGCCCCAACAGCGAGGCCACCCACCTACACGAGCACATCACCAGCGGTGGCC 5760  
 Qy 5761 TCCAGCCAGCGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5820  
 Db 5761 TCCAGCCAGCGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5820  
 Qy 5821 TACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5880  
 Db 5821 TACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5880  
 Qy 5881 GAGAAGTCCCGCCGAGTGTCTCAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5940  
 Db 5881 GAGAAGTCCCGCCGAGTGTCTCAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5940  
 Qy 5941 GAAAGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6000  
 Db 5941 GAAAGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6000  
 Qy 6001 AACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6060  
 Db 6001 AACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6060  
 Qy 6061 GACTGGAATAAAAAGTGGACCACTGCCAGCTAACTGGTCACTCTGAGGATCATCCAG 6120  
 Db 6061 GACTGGAATAAAAAGTGGACCACTGCCAGCTAACTGGTCACTCTGAGGATCATCCAG 6120  
 Qy 6121 CTGAATGGAGAAATCCGGCAGCAGGTTGAAAAGTCTGTTCTGAGAACAGATTTATGCTGA 6180  
 Db 6121 CTGAATGGAGAAATCCGGCAGCAGGTTGAAAAGTCTGTTCTGAGAACAGATTTATGCTGA 6180  
 Qy 6181 GCAGGTTCAATGTGACTTCTAGACGTTGACTTAAAAAATGGCCCTTAAGGCTGCAGAGC 6240  
 Db 6181 GCAGGTTCAATGTGACTTCTAGACGTTGACTTAAAAAATGGCCCTTAAGGCTGCAGAGC 6240  
 Qy 6241 CAGCCACTCTGCTTACAAAAGAGTACTTAGTGCACTGACTGTAGTGAACAAATTTGTA 6300  
 Db 6241 CAGCCACTCTGCTTACAAAAGAGTACTTAGTGCACTGACTGTAGTGAACAAATTTGTA 6300

QY 5301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCC 6360  
Db 6301 AACCTCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCC 6360  
QY 6361 GAGAGACAAATCTCTGTGTTGATTTTGAAGGACAGGCAAGACCAACACTGTATTTAGTTCCA 6420  
Db 6361 GAGAGACAAATCTCTGTGTTGATTTTGAAGGACAGGCAAGACCAACACTGTATTTAGTTCCA 6420  
QY 6421 TAGCCAGGCTCTAAGAGGACAAAGTGGCTGGCCCTTAAAGAACACACAGATGACTGGAATG 6480  
Db 6421 TAGCCAGGCTCTAAGAGGACAAAGTGGCTGGCCCTTAAAGAACACACAGATGACTGGAATG 6480  
QY 6481 ATGTGTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGGAGTTAGCAATTCATTT 6540  
Db 6481 ATGTGTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGGAGTTAGCAATTCATTT 6540  
QY 6541 TTGGCTTAAGAAATCGAGAAATGTAGTTTGA 6574  
Db 6541 TTGGCTTAAGAAATCGAGAAATGTAGTTTGA 6574

RESULT 3  
US-10-325-430-11  
; Sequence 11, Application US/10325430  
; Publication No. US20030153525A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc  
; APPLICANT: Silos-Santiago, Imaculada  
; APPLICANT: Rosenfeld, Julie Beth  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING  
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,  
; FILE REFERENCE: 32838,336 and 52908  
; CURRENT APPLICATION NUMBER: US/10/325,430  
; PRIOR APPLICATION NUMBER: US 60/341,953  
; PRIOR FILING DATE: 2001-12-19  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 6162  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(6162)

US-10-325-430-11

Query Match 93.7%; Score 6162; DB 15; Length 6162;  
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;  
Matches 6162; Conservative 0; Mismatches 0

QY 19 ATGTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGTGCTCTGAACCCAT 78  
Db 1 ATGTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGTGCTCTGAACCCAT 60

QY 79 GCAGCGGGCTCCAGGCTGAATCTGTTCCAGGGGAAACCAACCTTTATGACTCAA 138  
Db 61 GCAGCGGGCTCCAGGCTGAATCTGTTCCAGGGGAAACCAACCTTTATGACTCAA 120

QY 139 CAGCAGATGTCCTCTTTCCAGAGAGGATATTAGATGCCCTCTTTGCTCTCTTTGAA 198  
Db 121 CAGCAGATGTCCTCTTTCCAGAGAGGATATTAGATGCCCTCTTTGCTCTCTTTGAA 180

QY 199 GAATGAGTCAGCTGCTGTGATGAAGATTAGCAGCTGAGCAACTTTGTCGGAGAT 258  
Db 181 GAATGAGTCAGCTGCTGTGATGAAGATTAGCAGCTGAGCAACTTTGTCGGAGAT 240

QY 259 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCCTTCGGCAAGAGACTTCGAAGTCAGA 318  
Db 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCCTTCGGCAAGAGACTTCGAAGTCAGA 300

QY 319 AGTCTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAGAAAGCAACCGG 378  
Db 301 AGTCTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAGAAAGCAACCGG 360  
QY 379 GACATCTATGCTATGAAGTGTATGAAGAAGAGCTTTATTTGGCCAGGAGAGAGTTTCA 438  
Db 361 GACATCTATGCTATGAAGTGTATGAAGAAGAGCTTTATTTGGCCAGGAGAGAGTTTCA 420  
QY 439 TTTTTCGAGAGAGCGGAAACATATATCTCGAAGCAACAAGCCCGTGGATCCCCCAATTA 498  
Db 421 TTTTTCGAGAGAGCGGAAACATATATCTCGAAGCAACAAGCCCGTGGATCCCCCAATTA 480  
QY 499 CAGTATGCCCTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 558  
Db 481 CAGTATGCCCTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540  
QY 559 GACTTGTGTCACCTTTTGAATAGATATGAGGACAGTATGATGAAACCTGATACAGTTT 618  
Db 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACAGTATGATGAAACCTGATACAGTTT 600  
QY 619 TACCTAGCTAGCTGATTTTGGCTGTTTCACAGCGTTTCATCTGATGGGATACGTGCATCGA 678  
Db 601 TACCTAGCTAGCTGATTTTGGCTGTTTCACAGCGTTTCATCTGATGGGATACGTGCATCGA 660  
QY 679 GACATCAAGCTGAGAACATTCGTTGACGGACAGGACACATCANGCTGGTGGATTTT 738  
Db 661 GACATCAAGCTGAGAACATTCGTTGACGGACAGGACACATCANGCTGGTGGATTTT 720  
QY 739 GGATCTGCGCGAAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTTGGAGC 798  
Db 721 GGATCTGCGCGAAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTTGGAGC 780  
QY 799 CCAGATTACATGGCTCTGTAAGTGTCTGATGTAACCGGGGATGGAAGGACCTTAC 858  
Db 781 CCAGATTACATGGCTCTGTAAGTGTCTGATGTAACCGGGGATGGAAGGACCTTAC 840  
QY 859 GGCTCGACTGTGACTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 918  
Db 841 GGCTCGACTGTGACTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900  
QY 919 TCCCTCTTCGAGAGGAACTCTGCCAGAACCTTCAATAACATTAATTAATTTCCAGCGG 978  
Db 901 TCCCTCTTCGAGAGGAACTCTGCCAGAACCTTCAATAACATTAATTAATTTCCAGCGG 960  
QY 979 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1038  
Db 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020  
QY 1039 TTGTTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1098  
Db 1021 TTGTTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1080  
QY 1099 TCTAAAATTGACTGGAAACAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1158  
Db 1081 TCTAAAATTGACTGGAAACAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140  
QY 1159 TCTCAGCATGACACTCCAAATTTTGAAGAACAGAGAAATTCGTGGGTTCATCTCTCT 1218  
Db 1141 TCTCAGCATGACACTCCAAATTTTGAAGAACAGAGAAATTCGTGGGTTCATCTCTCT 1200  
QY 1219 CCCTGCGAGCTGAGCCCTCAGGCTTCGCGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1278  
Db 1201 CCCTGCGAGCTGAGCCCTCAGGCTTCGCGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1260  
QY 1279 TACAGAGGACACTGGGATTTCTTGGTAGATCTGAGTCTGTTGTGCGGTCTGGACTCC 1338  
Db 1261 TACAGAGGACACTGGGATTTCTTGGTAGATCTGAGTCTGTTGTGCGGTCTGGACTCC 1320  
QY 1339 CCTCCCAAGACTAGCTCCATGGAAAGAAACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1398  
Db 1321 CCTCCCAAGACTAGCTCCATGGAAAGAAACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380  
QY 1399 TCTCAGGACAAAGTGTCAAGATGGAGCAAGAAATGACCCCGTTTACATCGGAGAGTCTCA 1458

Db 1381 TCTCAGGACAAAGTGTACAAAGATGAGCAGGAGAAATGACCCGGTTACATCGGAGAGTGCA 1440  
QY 1459 GAGGTGGAGCTTGCTTAGTACAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518  
Db 1441 GAGGTGGAGGCTGTGCTTAGTACAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
QY 1519 TCCCTCCTGGAGCAGGACCTTCTACCTACCTACATCAGAGATGAGTAGCTTAAAGCGAAGT 1578  
Db 1501 TCCCTCCTGGAGCAGGACCTTCTACCTACCTACATCAGAGATGAGTAGCTTAAAGCGAAGT 1560  
QY 1579 TTGGAGCAAGCAGGATGAGGTGTCACAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1638  
Db 1561 TTGGAGCAAGCAGGATGAGGTGTCACAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620  
QY 1639 GATATCAGAGACGAGCGGAGGCTTCCAAAGAAATCAAAGAGCAGAGGTACAGGCTCAA 1698  
Db 1621 GATATCAGAGACGAGCGGAGGCTTCCAAAGAAATCAAAGAGCAGAGGTACAGGCTCAA 1680  
QY 1699 GTGGAAGAAATGAGTTGATGATGAATCAGTTTGGAGGATGTTGTCTCAGCAAGAAGA 1758  
Db 1681 GTGGAAGAAATGAGTTGATGATGAATCAGTTTGGAGGATGTTGTCTCAGCAAGAAGA 1740  
QY 1759 CGAGTGATCTCTAGAACTGAGCTGAGAGTCTCGGCTGCTGCTGAAGATTCAG 1818  
Db 1741 CGAGTGATCTCTAGAACTGAGCTGAGAGTCTCGGCTGCTGCTGAAGATTCAG 1800  
QY 1819 CGGAAAGCAGAGATGTCAGATATACTGTTTGAAGGCTTAAGGATCAAGGGAAGGCTGAA 1878  
Db 1801 CGGAAAGCAGAGATGTCAGATATACTGTTTGAAGGCTTAAGGATCAAGGGAAGGCTGAA 1860  
QY 1879 GTGGAGATATGCGAAATGAGAGATCAATGCTGAGAGCAGCTCAAAATTCAGAG 1938  
Db 1861 GTGGAGATATGCGAAATGAGAGATCAATGCTGAGAGCAGCTCAAAATTCAGAG 1920  
QY 1939 CTCGAAAGAACTCGAGAGGCTCCAAAGGAGCGAGCGGAGGAGCTGAGAGAGCTG 1998  
Db 1921 CTCGAAAGAACTCGAGAGGCTCCAAAGGAGCGAGCGGAGGAGCTGAGAGAGCTG 1980  
QY 1999 CAGAACCGAGAGATTCCTTGAAGGATCAAGAAAGAGCTGGTGAAGCTGAGGAAGC 2058  
Db 1981 CAGAACCGAGAGATTCCTTGAAGGATCAAGAAAGAGCTGGTGAAGCTGAGGAAGC 2040  
QY 2059 CGCCATCTCTGGAGAACAGGTAAAGAGACTAGAGCCATGAGGCGTAGAGAAACAGA 2118  
Db 2041 CGCCATCTCTGGAGAACAGGTAAAGAGACTAGAGCCATGAGGCGTAGAGAAACAGA 2100  
QY 2119 CTGAAGGATGACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCG 2178  
Db 2101 CTGAAGGATGACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCG 2160  
QY 2179 GAGCTCGAGAGAAACATCGGAGGCCCAAGTCTCAGCCGACACCTAGAAAGTGCACTG 2238  
Db 2161 GAGCTCGAGAGAAACATCGGAGGCCCAAGTCTCAGCCGACACCTAGAAAGTGCACTG 2220  
QY 2239 AAAACAGAAAGACGACACTATGAGGAAAGATTAAGTGTGGCAATCAGATAAGAAA 2298  
Db 2221 AAACAGAAAGACGACACTATGAGGAAAGATTAAGTGTGGCAATCAGATAAGAAA 2280  
QY 2299 GACCTGGTGCAAGAGACACTGAGAACATGATGACAGACACGAGGAGGAGGCCAT 2358  
Db 2281 GACCTGGTGCAAGAGACACTGAGAACATGATGACAGACACGAGGAGGAGGCCAT 2340  
QY 2359 GAGAGGCGCAAAATCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCGAAGATC 2418  
Db 2341 GAGAGGCGCAAAATCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCGAAGATC 2400  
QY 2419 AGATCCCTGGAAACAGAGGATTTGGAATGCTGTAAGCCCAATAAACTTGCAGCAAAATAGC 2478  
Db 2401 AGATCCCTGGAAACAGAGGATTTGGAATGCTGTAAGCCCAATAAACTTGCAGCAAAATAGC 2460  
QY 2479 AGTCTTTTACCCAAAGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2538

Db 2461 AGTCTTTTACCCAAAGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520  
QY 2539 CAGNAAATTTTACCTGGAGACACAGGCTGGAAAGTTGGAGGCCCAGAACCGAAACTGGAG 2598  
Db 2521 CAGNAAATTTTACCTGGAGACACAGGCTGGAAAGTTGGAGGCCCAGAACCGAAACTGGAG 2580  
QY 2599 GAGCAGCTGAGAGAGATCAGCCACCAAGACCAAGTGAACAAGAAATCGGCTGCTGGAACTG 2658  
Db 2581 GAGCAGCTGAGAGAGATCAGCCACCAAGACCAAGTGAACAAGAAATCGGCTGCTGGAACTG 2640  
QY 2659 GAGCAAGATTTGGGAGGTGAGTCTAGAGACAGAGAGCAGAAAATGAGAGCTCAAGCGC 2718  
Db 2641 GAGCAAGATTTGGGAGGTGAGTCTAGAGACAGAGAGCAGAAAATGAGAGCTCAAGCGC 2700  
QY 2719 CAGCTCACAGAGTACAGCTCTCCCTGCAGAGCGGAGTACAGTTGACAGCCCTGCGAG 2778  
Db 2701 CAGCTCACAGAGTACAGCTCTCCCTGCAGAGCGGAGTACAGTTGACAGCCCTGCGAG 2760  
QY 2779 GTTGACCGGCGGCTTGGAGAGCCAGCTTTCGCGAGCGAAGACAGAGCTTGGAAAGAGACC 2838  
Db 2761 GCTGCAAGGCGGCTTGGAGAGCCAGCTTTCGCGAGCGAAGACAGAGCTTGGAAAGAGACC 2820  
QY 2839 ACAGCAGAGCTGAAGAGGAGATCCAGGCACTCAGGCACTACGSCACATAGAGATGAATCCAGCGC 2898  
Db 2821 ACAGCAGAGCTGAAGAGGAGATCCAGGCACTCAGGCACTACGSCACATAGAGATGAATCCAGCGC 2880  
QY 2899 AAATTTGATCTCTTCGTAAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC 2958  
Db 2881 AAATTTGATCTCTTCGTAAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC 2940  
QY 2959 CAGCTGACCGAGACACGCTGAACTCAACAAACAAACTTCTACTTGTCCAAACACTC 3018  
Db 2941 CAGCTGACCGAGACACGCTGAACTCAACAAACAAACTTCTACTTGTCCAAACACTC 3000  
QY 3019 GATGAGGCTTCTCGCGCAACGACGAGATGTACAACTGCGAAAGTGAAGTGGACCATCTC 3078  
Db 3001 GATGAGGCTTCTCGCGCAACGACGAGATGTACAACTGCGAAAGTGAAGTGGACCATCTC 3060  
QY 3079 CGCGGAGATACCGAACGAGAGATGACGCTTACAGCCAGAAAGCAAAACGATGGAGGCT 3138  
Db 3061 CGCGGAGATACCGAACGAGAGATGACGCTTACAGCCAGAAAGCAAAACGATGGAGGCT 3120  
QY 3139 CTGAAGACACGCTGACCATGCTGGAGGAAACAGTCAATGGATTTGGAGGCCCTTAAACGAT 3198  
Db 3121 CTGAAGACACGCTGACCATGCTGGAGGAAACAGTCAATGGATTTGGAGGCCCTTAAACGAT 3180  
QY 3199 GAGCTGCTAGAAAAAGAGCGGAGTGGAGGCTTGGAGGAGCGCTCTGGGTGATGAGAAA 3258  
Db 3181 GAGCTGCTAGAAAAAGAGCGGAGTGGAGGCTTGGAGGAGCGCTCTGGGTGATGAGAAA 3240  
QY 3259 TCCAGTTTGAAGTGTGGGTTGAGAGCTGAGAGATGCTGCACACCGAGAAACAGAGC 3318  
Db 3241 TCCAGTTTGAAGTGTGGGTTGAGAGCTGAGAGATGCTGCACACCGAGAAACAGAGC 3300  
QY 3319 AGGCGAGAGCCGATCAGCGGATCAGCGAGTCTCGCGAGTGGTGGAGCTGGCAGTGAAG 3378  
Db 3301 AGGCGAGAGCCGATCAGCGGATCAGCGAGTCTCGCGAGTGGTGGAGCTGGCAGTGAAG 3360  
QY 3379 GAGCACAAGGCTGAGATTTCTGCTGCGAGCGGCTCTCAAAGAGCAGAGCTGAAAGGCC 3438  
Db 3361 GAGCACAAGGCTGAGATTTCTGCTGCTGCGAGCGGCTCTCAAAGAGCAGAGCTGAAAGGCC 3420  
QY 3439 GAGAGCTCTCTCACAAGCTCAATGATGCTGGAGAAAGCATGCTATGCTTGAATGAAT 3498  
Db 3421 GAGAGCTCTCTCACAAGCTCAATGATGCTGGAGAAAGCATGCTATGCTTGAATGAAT 3480  
QY 3499 GCCCGAAGCTTACAGCAGAGCTGGAGCTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3558  
Db 3481 GCCCGAAGCTTACAGCAGAGCTGGAGCTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3540  
QY 3559 GAGCAGCCAAATACAGCAGAGCTGGAGCTGCAAAAAATCACAATTTCCGCTGACT 3618  
Db 3541 GAGCAGCCAAATACAGCAGAGCTGGAGCTGCAAAAAATCACAATTTCCGCTGACT 3600

QY	3619	CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGA	CTTG	3678
Db	3601	CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGA	CTTG	3660
QY	3679	GAGTATCAGCTGGAAACATTTCAAGTTCTCTATTCTCATGAAGTGA	AAATGGAAGGC	3738
Db	3661	GAGTATCAGCTGGAAACATTTCAAGTTCTCTATTCTCATGAAGTGA	AAATGGAAGGC	3720
QY	3739	ACTATTTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGC	CAAAATGGACCAACTGCT	3798
Db	3721	ACTATTTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGC	CAAAATGGACCAACTGCT	3780
QY	3799	AAAGAAGAAAAGGGTTTATTATTGTCAGCGAAGAGGACCTCTGTTTAC	CCACACAGGTT	3858
Db	3781	AAAGAAGAAAAGGGTTTATTATTGTCAGCGAAGAGGACCTCTGTTTAC	CCACACAGGTT	3840
QY	3859	CCCTGTCAGTACAATGAGCTGAAGCTGGCCCTGGAGAAGGAGAAAGCT	CGCTGTGCAGAG	3918
Db	3841	CCCTGTCAGTACAATGAGCTGAAGCTGGCCCTGGAGAAGGAGAAAGCT	CGCTGTGCAGAG	3900
QY	3919	CTAGAGGAAGCCCTTTCAGAAAGCCGCGATCGAGCTCCGTC	CCCGGAGGAAGCTGCC	3978
Db	3901	CTAGAGGAAGCCCTTTCAGAAAGCCGCGATCGAGCTCCGTC	CCCGGAGGAAGCTGCC	3960
QY	3979	CACCGAAGACGAGGACACCCACACACCCATCCACGCCAGCACCCG	GAGGACAGATC	4038
Db	3961	CACCGAAGACGAGGACACCCACACACCCATCCACGCCAGCACCCG	GAGGACAGATC	4020
QY	4039	GCCATGTCGCCATCGTCGGTCCGACAGCACAGCCAGTGC	CACTGACCTGTGCGC	4098
Db	4021	GCCATGTCGCCATCGTCGGTCCGACAGCACAGCCAGTGC	CACTGACCTGTGCGC	4080
QY	4099	CCGCCATCCAGCCGACGAAAGGAGTCTTAACTCCAGAGGAATTTAGT	CGCGCTCTTAAG	4158
Db	4081	CCGCCATCCAGCCGACGAAAGGAGTCTTAACTCCAGAGGAATTTAGT	CGCGCTCTTAAG	4140
QY	4159	GAACGCAATGACACACAATATTCCTCACCGATTC	CAACGTAGACTGAACATGCGAGCCACA	4218
Db	4141	GAACGCAATGACACACAATATTCCTCACCGATTC	CAACGTAGACTGAACATGCGAGCCACA	4200
QY	4219	AAAGTGCTGTGTCTGATACCGTGCATTTTGGACCGCAGGCAT	CCAAAATGCTCGAA	4278
Db	4201	AAAGTGCTGTGTCTGATACCGTGCATTTTGGACCGCAGGCAT	CCAAAATGCTCGAA	4260
QY	4279	TGTCAGGTGATGTGTACCCCAAGTGTCCACGTGTTCCACGAC	CACTGCGGCTTGCT	4338
Db	4261	TGTCAGGTGATGTGTACCCCAAGTGTCCACGTGTTCCACGAC	CACTGCGGCTTGCT	4320
QY	4339	GCTGAATATGCCACACATTTACCGAGGCTTCTGCGCTGAC	AAAAATGAATCCCAAGT	4398
Db	4321	GCTGAATATGCCACACATTTACCGAGGCTTCTGCGCTGAC	AAAAATGAATCCCAAGT	4380
QY	4399	CTTCAGACCAAGAGGCCACGACAGCTTGCACTCGAAGGGT	GTGATGAAGTGCCAGG	4458
Db	4381	CTTCAGACCAAGAGGCCACGACAGCTTGCACTCGAAGGGT	GTGATGAAGTGCCAGG	4440
QY	4459	AAATAACAAACGAGCAGCAAGGCTGGACAGGAAGTACATCTCT	GAGGGATCAAAA	4518
Db	4441	AAATAACAAACGAGCAGCAAGGCTGGACAGGAAGTACATCTCT	GAGGGATCAAAA	4500
QY	4519	GTGCTCATTTATGACAAATGAAGCCAGAGAGCTGGACAGAG	CGCGTGGAGAAATTTGAG	4578
Db	4501	GTGCTCATTTATGACAAATGAAGCCAGAGAGCTGGACAGAG	CGCGTGGAGAAATTTGAG	4560
QY	4579	CTGTGCTTCCGACCGGGATGTATCTATCATGTTGCGGTG	GTGCTTCCGAACTCGCA	4638
Db	4561	CTGTGCTTCCGACCGGGATGTATCTATCATGTTGCGGTG	GTGCTTCCGAACTCGCA	4620
QY	4639	AATACAGCCAAAGCAGAAAAAGCAGAGCTGATGCTAAAC	TGCTTGGAACTCCCTGCTG	4698
Db	4621	AATACAGCCAAAGCAGAAAAAGCAGAGCTGATGCTAAAC	TGCTTGGAACTCCCTGCTG	4680

QY	4699	AAACTGGAAGGTGATGACCGGTCTTAGACATGTAAGCTGACGCGCTGCCCTTCAGTGACCAAGGTG	4750
DB	4681	AAACTGGAAGGTGATGACCGGTCTTAGACATGAACTGCACGCTGCGCTTCAGTGACCAAGGTG	4740
QY	4759	GTGTTGGTGGCCACCCAGGAAGGGCTCTACGCGCTGAAATGTCTTGAATAACTCCCTAAACC	4818
DB	4741	GTGTTGGTGGCCACCCAGGAAGGGCTCTACGCGCTGAAATGTCTTGAATAACTCCCTAAACC	4800
QY	4819	CATGTCCTCCAGGAATTCGAGCAGTCTTCCAAAATTTATATATTAAAGACCTCGAGAAGCTA	4878
DB	4801	CATGTCCTCCAGGAATTCGAGCAGTCTTCCAAAATTTATATATTAAAGACCTCGAGAAGCTA	4860
QY	4879	CTCATGATAGCAGGAAGAGCGGGCACTGTGCTTGTGTGACGCTGAAGAAAGTGAACACAG	4938
DB	4861	CTCATGATAGCAGGAAGAGCGGGCACTGTGCTTGTGTGACGCTGAAGAAAGTGAACACAG	4920
QY	4939	TCCTTGGCCCGAGTCCCAACCTGCTGCCCCAGCCGACATCTCACCCAACTTTTGTGAAGCT	4998
DB	4921	TCCTTGGCCCGAGTCCCAACCTGCTGCCCCAGCCGACATCTCACCCAACTTTTGTGAAGCT	4980
QY	4999	GTCAAGGGCTGCCACTTGTGTTGGGGCAGGCAGATTTGAAACGGGCTCTGCATCTGTGCA	5058
DB	4981	GTCAAGGGCTGCCACTTGTGTTGGGGCAGGCAGATTTGAAACGGGCTCTGCATCTGTGCA	5040
QY	5059	GCCATGCCCAGCAAAAGTCGTCAATTTCTCCGCTACAACGAAAACTCTACGAAATACTGCATC	5118
DB	5041	GCCATGCCCAGCAAAAGTCGTCAATTTCTCCGCTACAACGAAAACTCTACGAAATACTGCATC	5100
QY	5119	CGAAAGAGATAGACCTCAGAGCCCTGAGAGCTGTATCCACTTACACCAATTCAGTATC	5178
DB	5101	CGAAAGAGATAGACCTCAGAGCCCTGAGAGCTGTATCCACTTACACCAATTCAGTATC	5160
QY	5179	CTCATTTGGAAACCAATAAATTTCTACGAAATTCGACATGAAGCAGTACACGCTCGAGGAATTC	5238
DB	5161	CTCATTTGGAAACCAATAAATTTCTACGAAATTCGACATGAAGCAGTACACGCTCGAGGAATTC	5220
QY	5239	CTGGATAAGAATGACCAATTCCTTGGCACCTGTGTGTGTTGCCGCTCTTCCACACGCTTC	5298
DB	5221	CTGGATAAGAATGACCAATTCCTTGGCACCTGTGTGTGTTGCCGCTCTTCCACACGCTTC	5280
QY	5299	CCTGTCTCAATCGTCGAGGTGAACACGGCAGGCGCAGGAGGAGTACTTGCTGTGTTTC	5358
DB	5281	CCTGTCTCAATCGTCGAGGTGAACACGGCAGGCGCAGGAGGAGTACTTGCTGTGTTTC	5340
QY	5359	CACGAATTTGGAGTGTCTGTGGATTTCTACGGAAGACGTAGCCGCGACAGACGATCTCAAG	5418
DB	5341	CACGAATTTGGAGTGTCTGTGGATTTCTACGGAAGACGTAGCCGCGACAGACGATCTCAAG	5400
QY	5419	TGAGAGTCGCTTACCTTTGGCCTTGGCTACAGAGAACCTCTATCTGTTGTGACCCACTTC	5478
DB	5401	TGAGAGTCGCTTACCTTTGGCCTTGGCTACAGAGAACCTCTATCTGTTGTGACCCACTTC	5460
QY	5479	AACCTCACTCGAAGTAATTGAGATCCAGGACACGCTCTCTCAGCAGGAGCCCTGTGCCCGACG	5538
DB	5461	AACCTCACTCGAAGTAATTGAGATCCAGGACACGCTCTCTCAGCAGGAGCCCTGTGCCCGACG	5520
QY	5539	TACCTGGACATCCGAAACCCGCGCTACCTTGGGCCCTTGCAATTTCTCAGAGGAGGATTTAC	5598
DB	5521	TACCTGGACATCCGAAACCCGCGCTACCTTGGGCCCTTGCAATTTCTCAGAGGAGGATTTAC	5580
QY	5599	TTGGCGTCTCATACACAGGATAAATTAAGGTCATTTGCTGCAAGGGAACCTCTGCTGAAG	5658
DB	5581	TTGGCGTCTCATACACAGGATAAATTAAGGTCATTTGCTGCAAGGGAACCTCTGCTGAAG	5640
QY	5659	GAGTCCGGCACTGAACACCAACCGGGGCCGTCCACCTCCCGCAGCAGCCCCCAACAGCGCA	5718
DB	5641	GAGTCCGGCACTGAACACCAACCGGGGCCGTCCACCTCCCGCAGCAGCCCCCAACAGCGCA	5700
QY	5719	GGCCCAACCACTGACAAACGAGACATCAACAGCGCTGCCTCCAGCCACGCGCCGCC	5778
DB	5701	GGCCCAACCACTGACAAACGAGACATCAACAGCGCTGCCTCCAGCCACGCGCCGCC	5760
QY	5779	GAAGGCCCCACGACCCGGCAGAGCAACGACACCCACCGCTTACCGGAGGGCGGAC	5838



1381 TCTCAGGACAAGTGTACACAGATGGAGCAGGAATAATGACCGGTATACATCGGAGAGTGTCA 1440  
1459 GAGGTGGAGGCTGTGTTAGTCAGAAAGGAGGTGGAGCTGAAGGCCCTCTCAGACTCAGAGA 1518  
1441 GAGGTGGAGGCTGTGTTAGTCAGAAAGGAGGTGGAGCTGAAGGCCCTCTCAGACTCAGAGA 1500  
1519 TCCTCTCTGGAGCAGGACCTTGCTACTACATACAGAAATGAGTACAGTCTGAAGGCAAGT 1578  
1501 TCCTCTCTGGAGCAGGACCTTGCTACTACATACAGAAATGAGTACAGTCTGAAGGCAAGT 1560  
1579 TTGGAGCAAGCAGGATGAGGTTGCCAGGAGGATGCAAGAGCACTGCAGCTTCTCCAT 1638  
1561 TTGGAGCAAGCAGGATGAGGTTGCCAGGAGGATGCAAGAGCACTGCAGCTTCTCCAT 1620  
1639 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGACAGGAGTACCAGGCTCAA 1698  
1621 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGACAGGAGTACCAGGCTCAA 1680  
1699 GTGGAAGAAATGAGTGTGATGATGATCAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1758  
1681 GTGGAAGAAATGAGTGTGATGATGATGATCAATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740  
1759 CGGAGTGATCTCTACGAATCTCAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1818  
1741 CGGAGTGATCTCTACGAATCTCAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1800  
1819 CGGAAGCAGACAAATGTGAGCAATAAATCTGTAAGGCTTAAGGATCAAGGAGGCTGAA 1878  
1801 CGGAAGCAGACAAATGTGAGCAATAAATCTGTAAGGCTTAAGGATCAAGGAGGCTGAA 1860  
1879 GTGGAGAGATATCGGAAATCTGAGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938  
1861 GTGGAGAGATATCGGAAATCTGAGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
1939 CTCGAAGAGAACTGGAGAGGCTGCAAGAGGAGCGAGCCGAGGAGGAGTGGAGAGGCTG 1998  
1921 CTCGAAGAGAACTGGAGAGGCTGCAAGAGGAGCGAGCCGAGGAGGAGTGGAGAGGCTG 1980  
1999 CAGAACCGAGAGATCTCTGAGGAGCATCAGAAAGAGCTGAGAGAGCTGAGAGAGCTG 2058  
1981 CAGAACCGAGAGATCTCTGAGGAGCATCAGAAAGAGCTGAGAGAGCTGAGAGAGCTG 2040  
2059 CGCATTTCTCTGAGAGAACTGAGAGAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGA 2118  
2041 CGCATTTCTCTGAGAGAACTGAGAGAGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGA 2100  
2119 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATAAATTTCTG 2178  
2101 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATAAATTTCTG 2160  
2179 GAGCTCGAAGAGAAAATCGGAGGAGGAGGAGTCTCAGCCAGCAGCTAGAGAGTGCACCTG 2238  
2161 GAGCTCGAAGAGAAAATCGGAGGAGGAGGAGTCTCAGCCAGCAGCTAGAGAGTGCACCTG 2220  
2239 AAACAGAAAGAGCAGCCTATGAGGAAAGATTAAGTTGTTGGACAATCAGATAAGAAA 2298  
2221 AAACAGAAAGAGCAGCCTATGAGGAAAGATTAAGTTGTTGGACAATCAGATAAGAAA 2280  
2299 GAGCTGGCTGACAAGGAGACATGGAGAACATGATGAGAGACAGAGAGGAGGAGGAGGAGG 2358  
2281 GAGCTGGCTGACAAGGAGACATGGAGAACATGATGAGAGACAGAGAGGAGGAGGAGGAGG 2340  
2359 GAGAGGGCAAAATTTCTAGCGAAACAGAGGAGGAGTCAATGCTATGATGATTTCCAGATC 2418  
2341 GAGAGGGCAAAATTTCTAGCGAAACAGAGGAGGAGTCAATGCTATGATGATTTCCAGATC 2400  
2419 AGATCCCTGGACAGAGGATTTGGAACTGCTGAGGCCAATAAATTTGCGAGCAATAGC 2478  
2401 AGATCCCTGGACAGAGGATTTGGAACTGCTGAGGCCAATAAATTTGCGAGCAATAGC 2460  
2479 AGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAA 2538

2461 AGTCTTTTACCCAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520  
2539 CAGAAATTTTACCTGGAGACACAGGCTTGGAAAGTTGGAGCCCAAGAACCCGAAACTGGAG 2598  
2521 CAGAAATTTTACCTGGAGACACAGGCTTGGAAAGTTGGAGCCCAAGAACCCGAAACTGGAG 2580  
2599 GAGCAGCTGGAGAAAGATCAGCCACCAAGACACAGTGCACAAAGATCGGCTGCTGGAAGCTG 2658  
2581 GAGCAGCTGGAGAAAGATCAGCCACCAAGACACAGTGCACAAAGATCGGCTGCTGGAAGCTG 2640  
2659 GAGCAAGATTTGCGGAGGAGTCACTAGAGCAGCAGGAGCAGAAACTGGAGCTCAAGCGC 2718  
2641 GAGCAAGATTTGCGGAGGAGTCACTAGAGCAGCAGGAGCAGAAACTGGAGCTCAAGCGC 2700  
2719 CAGCTCAGAGCTACAGCTCTCTCTCAGAGGAGGAGTCAAGTTCAGAGCTCTGAG 2778  
2701 CAGCTCAGAGCTACAGCTCTCTCTCAGAGGAGGAGTCAAGTTCAGAGCTCTGAG 2760  
2779 GCTCAGCGGCGGCTCTGGAGAGCAGCTTTCGCGAGGAGGAGCAGAGCTGGAGAGACC 2838  
2761 GCTCAGCGGCGGCTCTGGAGAGCAGCTTTCGCGAGGAGGAGCAGAGCTGGAGAGACC 2820  
2839 ACAGCAGAACTGAAGAGGAGATCCAGGCACTCAAGCACAATAGAGATGAAATCCAGCGC 2898  
2821 ACAGCAGAACTGAAGAGGAGATCCAGGCACTCAAGCACAATAGAGATGAAATCCAGCGC 2880  
2899 AAATTTGATGCTCTTCTGTAACGCTGTACTGTAAATCAGACACCTGGAGGAGGAGCTAAAC 2958  
2881 AAATTTGATGCTCTTCTGTAACGCTGTACTGTAAATCAGACACCTGGAGGAGGAGCTAAAC 2940  
2959 CAGCTGACCGAGGACAACTGTAACCTCAACCAAACTTCTACTTTGCTCCAAACAACTC 3018  
2941 CAGCTGACCGAGGACAACTGTAACCTCAACCAAACTTCTACTTTGCTCCAAACAACTC 3000  
3019 GATGAGGCTTTCTGGCGCAACGAGAGATGTAACAATCGGAGTGAAGTGGACATCTC 3078  
3001 GATGAGGCTTTCTGGCGCAACGAGAGATGTAACAATCGGAGTGAAGTGGACATCTC 3060  
3079 CGCGGAGATCAACGAAACGAGAGATGCACTTACCAGCAGAAACCAAGCATGGAGGCT 3138  
3061 CGCGGAGATCAACGAAACGAGAGATGCACTTACCAGCAGAAACCAAGCATGGAGGCT 3120  
3139 CTGAAGACCACTGTCACCATGCTGGAGGAAACAGGTCAATGATTTGGAGGCTTAAACGAT 3198  
3121 CTGAAGACCACTGTCACCATGCTGGAGGAAACAGGTCAATGATTTGGAGGCTTAAACGAT 3180  
3199 GAGCTGTAGAAAAGAGCGGAGTGGAGGCTGGAGGAGGCTCTGGTGTATGAGAAA 3258  
3181 GAGCTGTAGAAAAGAGCGGAGTGGAGGCTGGAGGAGGCTCTGGTGTATGAGAAA 3240  
3259 TCCAGTTTGAAGTGTCCGTTCCAGAGCTGACAGAAATCTGGACACCGAGAAACAGAGC 3318  
3241 TCCAGTTTGAAGTGTCCGTTCCAGAGCTGACAGAAATCTGGACACCGAGAAACAGAGC 3300  
3319 AGGCGAGAGCCGATCAGCGGATCCCGAGTCTCGGAGGAGGCTGGAGGAGGCTGGAGGAG 3378  
3301 AGGCGAGAGCCGATCAGCGGATCCCGAGTCTCGGAGGAGGCTGGAGGAGGCTGGAGGAG 3360  
3379 GAGCAAGGCTGAGATTTCTCGCTCTGAGCAGGCTCTCAAGAGCAGAACTGTAAGGCTC 3438  
3361 GAGCAAGGCTGAGATTTCTCGCTCTGAGCAGGCTCTCAAGAGCAGAACTGTAAGGCTC 3420  
3439 GAGAGCTCTCTGACAGCTCAATGACCTGGAGAAAGCATGCTATGCTGTAATGAAT 3498  
3421 GAGAGCTCTCTGACAGCTCAATGACCTGGAGAAAGCATGCTATGCTGTAATGAAT 3480  
3499 GCCGAGGCTTACAGCAGAACTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3558  
3481 GCCGAGGCTTACAGCAGAACTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3540  
3559 GAGCAAGCAGAAATTAAGCAGCAGAGTGGAGCTGACAGAAATTAAGCTTCTGCTGACT 3618  
3541 GAGCAAGCAGAAATTAAGCAGCAGAGTGGAGCTGACAGAAATTAAGCTTCTGCTGACT 3600

QY 3619 CAAGGACTGCAAGAGCTCTAGATCGGCGTGTACTGAAACACAGAAAGAGTACTTG 3678  
Db 3601 CAAGGACTGCAAGAGCTCTAGATCGGCGTGTACTGAAACACAGAAAGAGTACTTG 3660  
QY 3679 GAGTATCAGCTGGAAAAACATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGC 3738  
Db 3661 GAGTATCAGCTGGAAAAACATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGC 3720  
QY 3739 ACTATTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGCCAAAATGGACCAACTGCT 3798  
Db 3721 ACTATTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGCCAAAATGGACCAACTGCT 3780  
QY 3799 AAAAAAAGGGGTTTATTTAGTCGACCGAAAGAGGACCTGCTTTACCCACACAGGTT 3858  
Db 3781 AAAAAAAGGGGTTTATTTAGTCGACCGAAAGAGGACCTGCTTTACCCACACAGGTT 3840  
QY 3859 CTTCTGCAGTCAATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTGCTGTGCGAGAG 3918  
Db 3841 CTTCTGCAGTCAATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTGCTGTGCGAGAG 3900  
QY 3919 CTAGAGGAAGCCCTTCAGAAAGCCCGCATCGAGCTCCGGTCCGCCGGGAGGAAGCTGCC 3978  
Db 3901 CTAGAGGAAGCCCTTCAGAAAGCCCGCATCGAGCTCCGGTCCGCCGGGAGGAAGCTGCC 3960  
QY 3979 CACCGCAAGCAACGGACCAACCAACCAATCCACGCCAGCCACCGCGAGGCGAGATC 4038  
Db 3961 CACCGCAAGCAACGGACCAACCAACCAATCCACGCCAGCCACCGCGAGGCGAGATC 4020  
QY 4039 GGCATGTCGCCCATCGTGGCTGCGCAGAGCACCGCCAGTCCGATGAGCTGCTGCC 4098  
Db 4021 GGCATGTCGCCCATCGTGGCTGCGCAGAGCACCGCCAGTCCGATGAGCTGCTGCC 4080  
QY 4099 CGCCATCCAGCGCGCAAGAGGAGTCTTCAACTCCAGAGGAATTTAGTGGCGTCTTAAG 4158  
Db 4081 CGCCATCCAGCGCGCAAGAGGAGTCTTCAACTCCAGAGGAATTTAGTGGCGTCTTAAG 4140  
QY 4159 GAACGATGCACACAAATATTCCTACCGATTCACGTAGAGTGAATGCGAGCCACA 4218  
Db 4141 GAACGATGCACACAAATATTCCTACCGATTCACGTAGAGTGAATGCGAGCCACA 4200  
QY 4219 AAGTGTCGTGTCTGATACCGGTGCACTTTGACGCGCAGGCAATCAAAATGCTCGAA 4278  
Db 4201 AAGTGTCGTGTCTGATACCGGTGCACTTTGACGCGCAGGCAATCAAAATGCTCGAA 4260  
QY 4279 TGTGAGTGATGTGTACCCCAAGTGTCTCAAGTGTGCGAGCAACCTGCGGCTTGCT 4338  
Db 4261 TGTGAGTGATGTGTACCCCAAGTGTCTCAAGTGTGCGAGCAACCTGCGGCTTGCT 4320  
QY 4339 GCTGAATATGCCACACACTTCACGAGGCTTCTCCGTGACAAATGAATCCCGAGGT 4398  
Db 4321 GCTGAATATGCCACACACTTCACGAGGCTTCTCCGTGACAAATGAATCCCGAGGT 4380  
QY 4399 CTCCAGACCAAGGAGCCAGCAGCAGTTCACCTGGAGGGTGGATGAAGTGCCGAGG 4458  
Db 4381 CTCCAGACCAAGGAGCCAGCAGCAGTTCACCTGGAGGGTGGATGAAGTGCCGAGG 4440  
QY 4459 AATACAAACGAGGACAGCAAGCTGGGACAGGAAGTACATTTGCTGGAGGGATCAAAA 4518  
Db 4441 AATACAAACGAGGACAGCAAGCTGGGACAGGAAGTACATTTGCTGGAGGGATCAAAA 4500  
QY 4519 GTCCCTCATTTATGCAATGAAGCCAGAGAAAGTGGACAGAGCCGGTGAAGAAATTTGAG 4578  
Db 4501 GTCCCTCATTTATGCAATGAAGCCAGAGAAAGTGGACAGAGCCGGTGAAGAAATTTGAG 4560  
QY 4579 CTGTGCTTCCCGACGGGATGTATCTATTATGTTGCGGTTGGTCTTCGAACTCGCA 4638  
Db 4561 CTGTGCTTCCCGACGGGATGTATCTATTATGTTGCGGTTGGTCTTCGAACTCGCA 4620  
QY 4639 AATACAGCCAAAGCAGAAAAGCAGAAAGTGTATTAATCTCTTGGAAACTCCCTGTG 4698  
Db 4621 AATACAGCCAAAGCAGAAAAGCAGAAAGTGTATTAATCTCTTGGAAACTCCCTGTG 4680

QY 4599 AAACCTGGAAGGTGATGATACCGTCTAGACATGAACCTGACCGCTGCCCTTCAGTGACCAAGTG 4758  
Db 4581 AAACCTGGAAGGTGATGATACCGTCTAGACATGAACCTGACCGCTGCCCTTCAGTGACCAAGTG 4740  
QY 4759 GTGTTGTGTGGCAACCGAGGAAGGCTCTACGCCCTGAATGCTTTGAAAAACTCCCTAAC 4818  
Db 4741 GTGTTGTGTGGCAACCGAGGAAGGCTCTACGCCCTGAATGCTTTGAAAAACTCCCTAAC 4800  
QY 4819 CATCTCCAGGAATTTGAGCAGCTTCCAAATTTATATTATCAAGGACCTGGAGAACTA 4878  
Db 4801 CATCTCCAGGAATTTGAGCAGCTTCCAAATTTATATTATCAAGGACCTGGAGAACTA 4860  
QY 4879 CTCATGATAGCAGGAGAAAGCGGCACTGTGCTTGTGGACGTTGAAGAAGTGAACAG 4938  
Db 4861 CTCATGATAGCAGGAGAAAGCGGCACTGTGCTTGTGGACGTTGAAGAAGTGAACAG 4920  
QY 4939 TCCCTGCGCAGTCCCACTTCCTGCCAGCCGACATCTCAACCAATTTTGAAGCT 4998  
Db 4921 TCCCTGCGCAGTCCCACTTCCTGCCAGCCGACATCTCAACCAATTTTGAAGCT 4980  
QY 4999 GTCAAGGCTGCCCACTTTGTTGGGGCAGGCAAGATTGAGAAACGGGCTCTGCATCTGTGCA 5058  
Db 4981 GTCAAGGCTGCCCACTTTGTTGGGGCAGGCAAGATTGAGAAACGGGCTCTGCATCTGTGCA 5040  
QY 5059 GCCATGCCAGCAAAAGTGTCTATTTCTCGGTCAACGAAAACTTCAGCAAAATACTGCATC 5118  
Db 5041 GCCATGCCAGCAAAAGTGTCTATTTCTCGGTCAACGAAAACTTCAGCAAAATACTGCATC 5100  
QY 5119 CGGAAAGAGATAGAGACTCAGAGCCCTGCAGCTGTATTCACCTTCACCAATTTACAGTATC 5178  
Db 5101 CGGAAAGAGATAGAGACTCAGAGCCCTGCAGCTGTATTCACCTTCACCAATTTACAGTATC 5160  
QY 5179 CTCATTGGAACCAATAATTTCTAGAAATCGAATGAAGCAGTACACGCTCGAGGAATTC 5238  
Db 5161 CTCATTGGAACCAATAATTTCTAGAAATCGAATGAAGCAGTACACGCTCGAGGAATTC 5220  
QY 5239 CTGATGAAGATGACCACTTCCTGGCACCTGCTGTGTTTCGCCGCTCTTCCACAGCTTC 5298  
Db 5221 CTGATGAAGATGACCACTTCCTGGCACCTGCTGTGTTTCGCCGCTCTTCCACAGCTTC 5280  
QY 5299 CCTGCTCAATTCGTGCAAGTGAACAGCGCAGGCGAGGAGGAGTACTTGTCTGTGTTTC 5358  
Db 5281 CCTGCTCAATTCGTGCAAGTGAACAGCGCAGGCGAGGAGGAGTACTTGTCTGTGTTTC 5340  
QY 5359 CACGAATTTGGAGTTGCTGGATTCTTACGGAAGAGCTAGCGCACAGAGCATCTCAAG 5418  
Db 5341 CACGAATTTGGAGTTGCTGGATTCTTACGGAAGAGCTAGCGCACAGAGCATCTCAAG 5400  
QY 5419 TGGAGTCTGCTTACCTTTGGCCTTTGCCCTACAGAGAACCTATCTGTTTGTGACCCACTTC 5478  
Db 5401 TGGAGTCTGCTTACCTTTGGCCTTTGCCCTACAGAGAACCTATCTGTTTGTGACCCACTTC 5460  
QY 5479 AACTCACTCGAAGTAAATTGAGATCCAGGCACGCTCTCTCAGAGGACCCCTGCCGAGCG 5538  
Db 5461 AACTCACTCGAAGTAAATTGAGATCCAGGCACGCTCTCTCAGAGGACCCCTGCCGAGCG 5520  
QY 5539 TACCTGGACATCCGGAACCCGCTACCTGGGCCCTGCCATTTCTCAGAGGCGATTTAC 5598  
Db 5521 TACCTGGACATCCGGAACCCGCTACCTGGGCCCTGCCATTTCTCAGAGGCGATTTAC 5580  
QY 5599 TTGCGCTCCTCATACCAAGGATAAATTAAAGGTCAATTTGCTGCAAGGGAACCTCGTGAAG 5658  
Db 5581 TTGCGCTCCTCATACCAAGGATAAATTAAAGGTCAATTTGCTGCAAGGGAACCTCGTGAAG 5640  
QY 5659 GAGTCCGGCATGNAACACACCGGGGCGCTCCACCTCCCGCAGCAGCCCCCAACAGGGA 5718  
Db 5641 GAGTCCGGCATGNAACACACCGGGGCGCTCCACCTCCCGCAGCAGCCCCCAACAGGGA 5700  
QY 5719 GGCCCAACCCAGTAAACACGAGCATCAACCAAGCGCTGCGCTCCAGCCCGAGCGCCGCC 5778  
Db 5701 GGCCCAACCCAGTAAACACGAGCATCAACCAAGCGCTGCGCTCCAGCCCGAGCGCCGCC 5760  
QY 5779 GAAGGCCCCAGCCACCCCGAGAGCCAAAGCACACCCCAACCGCTACCCCGAGGGCGGAGCC 5838





2893 QY CAGCGAAATTTGATGCTCTCTCTAACAAGCTGTACTGTAAATCAAGACCTGGAGGAGCAG 2952  
2977 Db CAGCGAAATTTGATGCTCTCTCTAACAAGCTGTACTGTAAATCAAGACCTGGAGGAGCAG 3036  
2953 QY CTAACACAGCTGACCGAGGACAAAGCTGTAACTCAACAAACCAAACTTCTACTTGTCCAAA 3012  
3037 Db CTAACACAGCTGACCGAGGACAAAGCTGTAACTCAACAAACCAAACTTCTACTTGTCCAAA 3096  
3013 QY CAATCGATGAGGCTTCTGCGCCCAACGACGAGATTGTACAACCTGCGAAAGTGAAGTGGAC 3072  
3097 Db CAATCGATGAGGCTTCTGCGCCCAACGACGAGATTGTACAACCTGCGAAAGTGAAGTGGAC 3156  
3073 QY CATCTCGCGCGGAGATCAAGGAAACGAGATGACGCTTACAGCCAGAACCAAGCATG 3132  
3157 Db CATCTCGCGCGGAGATCAAGGAAACGAGATGACGCTTACAGCCAGAACCAAGCATG 3216  
3133 QY GAGGCTCTGAAGACCACTGTCACCATGCTGGAGGAAACAGTCAATGATTGGAGGCCCTTA 3192  
3217 Db GAGGCTCTGAAGACCACTGTCACCATGCTGGAGGAAACAGTCAATGATTGGAGGCCCTTA 3276  
3193 QY AACGATGAGCTGTAGAAAAGAGCGGCAAGTGGAGGCCCTGGAGAGCGTCTCTGGGTGAT 3252  
3277 Db AACGATGAGCTGTAGAAAAGAGCGGCAAGTGGAGGCCCTGGAGAGCGTCTCTGGGTGAT 3336  
3253 QY GAGAAATCCAGTTTCTGAGTCTCGGTTTCGAGAGCTGCAGAGATGCTGGACACCCAGAAA 3312  
3337 Db GAGAAATCCAGTTTCTGAGTCTCGGTTTCGAGAGCTGCAGAGATGCTGGACACCCAGAAA 3396  
3313 QY CAGAGCGCGGAGAGCCGATCAGCGGATCAACGAGTCTCGCGAGTGTGGAGCTGGCA 3372  
3397 Db CAGAGCGCGGAGAGCCGATCAGCGGATCAACGAGTCTCGCGAGTGTGGAGCTGGCA 3456  
3373 QY GTGAAGGACACAGGCTGAGATTCTCGTCTGCAGCGGCTCTCAAGACGAGAGCTG 3432  
3457 Db GTGAAGGACACAGGCTGAGATTCTCGTCTGCAGCGGCTCTCAAGACGAGAGCTG 3516  
3433 QY AAGGCCGAGAGCCCTCTCTGACAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGA 3492  
3517 Db AAGGCCGAGAGCCCTCTCTGACAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGA 3576  
3493 QY ATCAATGCCCGAGCTTACAGAGAGCTGGAGCTGACAGAGCTCAACAGAGGCTT 3552  
3577 Db ATCAATGCCCGAGCTTACAGAGAGCTGGAGCTGACAGAGCTCAACAGAGGCTT 3636  
3553 QY CTGGAAGACCAAGCCAAATTAACAGAGAGATGGACCTGCAGAAAATCACATTTTCGT 3612  
3637 Db CTGGAAGACCAAGCCAAATTAACAGAGAGATGGACCTGCAGAAAATCACATTTTCGT 3696  
3613 QY CTGACTCAAGGACTGCAAGAGCTTAGATCGGGCTGTACTCTGAAAGACAGAAAGAT 3672  
3697 Db CTGACTCAAGGACTGCAAGAGCTTAGATCGGGCTGTACTCTGAAAGACAGAAAGAT 3756  
3673 QY GACTTGGAGTATCAGCTGGAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAATG 3732  
3757 Db GACTTGGAGTATCAGCTGGAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAATG 3816  
3733 QY GAAGGCACTATTCTCAACAAACCAAACTCATGATTTTCTGCAAGCCAAATGACCAA 3792  
3817 Db GAAGGCACTATTCTCAACAAACCAAACTCATGATTTTCTGCAAGCCAAATGACCAA 3876  
3793 QY CTTGCTAAAGAAAGAGGTTTATTAGTCAGGAAAGAGGACCTGCTTTACCCACA 3852  
3877 Db CTTGCTAAAGAAAGGTTTATTAGTCAGGAAAGAGGACCTGCTTTACCCACA 3892  
3853 QY CAGGTTCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAGGAAAGCTCGTGT 3912  
3893 Db AGGTTCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAGGAAAGCTCGTGT 3951  
3913 QY GCAGAGCTAGAGAGCCCTTCAAGAGCCGAGTCAGCTCCGTCGCCCGGAGGAA 3972  
3952 Db GCAGAGCTAGAGAGCCCTTCAAGAGCCGAGTCAGCTCCGTCGCCCGGAGGAA 4011  
3973 QY GCTGCCCGCAAGCAACGAGACCCACCCATCCAGCCAGCCACCGCGAGGAG 4032

4012 Db GCTGCCCGCAAGCAACGAGACCCACCCATCCACGCGAGCCACCGGAGCAG 4071  
4033 QY CAGATCGCATGTCCGCGCATCGTGGGTGCGAGACCCAGCCAGTGCATGAGCCTG 4092  
4072 Db CAGATCGCATGTCCGCGCATCGTGGGTGCGAGACCCAGCCAGTGCATGAGCCTG 4131  
4093 QY CTGCCCCGCGCATCCAGCGCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGCGCT 4152  
4132 Db CTGCCCCGCGCATCCAGCGCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGCGCT 4191  
4153 QY CTTAAGGAACGATGACCAACATATTCCTACCGATTCAACCTAGGACTGAACATGCGA 4212  
4192 Db CTTAAGGAACGATGACCAACATATTCCTACCGATTCAACCTAGGACTGAACATGCGA 4251  
4213 QY GCCACAAAGTGTGCTGTGTCTGGATACCGTGACCTTTGGACGCCAGGCATCCAAATGT 4272  
4252 Db GCCACAAAGTGTGCTGTGTCTGGATACCGTGACCTTTGGACGCCAGGCATCCAAATGT 4311  
4273 QY CTCGAATGTGAGTGATGTGTACCCCAAGTGTCCAGCTGCTTCCAGGCACCTGGGC 4332  
4312 Db CTCGAATGTGAGTGATGTGTACCCCAAGTGTCCAGCTGCTTCCAGGCACCTGGGC 4371  
4333 QY TTGCTGCTGAATATGCCACACACTTCCAGAGGCTTCTGCGGTGACAAAATGAATCC 4392  
4372 Db TTGCTGCTGAATATGCCACACACTTCCAGAGGCTTCTGCGGTGACAAAATGAATCC 4431  
4393 QY CCAGGCTCCAGACCAAGAGCCAGCAGCTTGCACCTGGAAGGTGGATGAAGTG 4452  
4432 Db CCAGGCTCCAGACCAAGAGCCAGCAGCTTGCACCTGGAAGGTGGATGAAGTG 4491  
4453 QY CCCAGGAATTAACAAACGAGACAGAGGCTGGACAGGAAGTACATTGTCTGGAGGA 4512  
4492 Db CCCAGGAATTAACAAACGAGACAGAGGCTGGACAGGAAGTACATTGTCTGGAGGA 4551  
4513 QY TCAAAAGTCTCTATTATGACAAATGAAGCCAGAGAGCTGGAAGGCGGTGGAGAA 4572  
4552 Db TCAAAAGTCTCTATTATGACAAATGAAGCCAGAGAGCTGGAAGGCGGTGGAGAA 4611  
4573 QY TTTGAGCTGTGCTTCCCGACGGGATGATCTATTATCATGCTGCTGCTTCCGAA 4632  
4612 Db TTTGAGCTGTGCTTCCCGACGGGATGATCTATTATCATGCTGCTGCTTCCGAA 4671  
4633 QY CTCGCAATACAGCCAAAGCA----- 4653  
4672 Db CTCGCAATACAGCCAAAGCAGATGCCATACATACTAGGAATGGAATCTCACCGCAC 4731  
4654 QY ----- 4654  
4732 Db ACCACTGTGCGCGGGAGAACCTCTACTTGTAGTCCAGCTTCCCTGTGACAAACAG 4791  
4654 QY -----GAAAAAGCA 4662  
4792 Db CGCTGGGTACCGCTTAGAATCAGTTGTCGAGGTGGAGAGTCTTAGGGAAAAAGCA 4851  
4663 QY GAAGCTGATGCTAAACTGTGTTGAAACTCTCCTGCTGAAACTGGAAGGTGATGACCGTCTA 4722  
4852 Db GAAGCTGATGCTAAACTGTGTTGAAACTCTCCTGCTGAAACTGGAAGGTGATGACCGTCTA 4911  
4723 QY GACATGAACCTGACGCTGCGCTTTCAGTGACAGGCTGCTGTTGGTGGCCACCGAGGAAGGG 4782  
4912 Db GACATGAACCTGACGCTGCGCTTTCAGTGACAGGCTGCTGTTGGTGGCCACCGAGGAAGGG 4971  
4783 QY CTCCTACGCCCTGAATGTCTTTGAAAAACTCTCCTAACCCATGTCCCGAAGATTTGAGCAGTC 4842  
4972 Db CTCCTACGCCCTGAATGTCTTTGAAAAACTCTCCTAACCCATGTCCCGAAGATTTGAGCAGTC 5031  
4843 QY TTCCAAATTTATATTAACAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGG 4902  
5032 Db TTCCAAATTTATATTAACAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGG 5091  
4903 QY GCACTGTGCTTGTGGACGTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTCGCTT 4962





QY 2791 GCCTTGAGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAGAGACACACAGAGCT 2850  
Db 2821 GCCTTGAGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAGAGACACACAGAGCT 2880  
QY 2851 GAAGAGGAGATCCAGGACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2910  
Db 2881 GAAGAGGAGATCCAGGACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2940  
QY 2911 CTTCTGAACAGCTGTACTGTATATCATCAGACCTGGAGGACAGCTAAACAGCTGACCGAG 2970  
Db 2941 CTTCTGAACAGCTGTACTGTATATCATCAGACCTGGAGGACAGCTAAACAGCTGACCGAG 3000  
QY 2971 GACAAAGCTGAATCAACCAACCAAACTTCTACTTTGTCTCAAAACAACCTCGATGAGGCTTCT 3030  
Db 3001 GACAAAGCTGAATCAACCAACCAAACTTCTACTTTGTCTCAAAACAACCTCGATGAGGCTTCT 3060  
QY 3031 GCGCCCAACGACGAGATTTGTACAACTGCGAAGTGAAGTGGACCATCTCCGCGGAGATC 3090  
Db 3061 GCGCCCAACGACGAGATTTGTACAACTGCGAAGTGAAGTGGACCATCTCCGCGGAGATC 3120  
QY 3091 ACGGAACGAGAGATGAGCTTACCAGCCAGAAGCAACGATGAGGCTCTGAAGACCAAG 3150  
Db 3121 ACGGAACGAGAGATGAGCTTACCAGCCAGAAGCAACGATGAGGCTCTGAAGACCAAG 3180  
QY 3151 TGCACCATCTGGAGGAACAGGTCTATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3210  
Db 3181 TGCACCATCTGGAGGAACAGGTCTATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240  
QY 3211 AAAGAGCGCAGTGGGAGCCCTGGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3270  
Db 3241 AAAGAGCGCAGTGGGAGCCCTGGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3300  
QY 3271 TGTGCGGTTTCGAGAGCTGCAGAGAAATGCTGACACCGAGAAACAGAGAGGCGAGAGCC 3330  
Db 3301 TGTGCGGTTTCGAGAGCTGCAGAGAAATGCTGACACCGAGAAACAGAGAGGCGAGAGCC 3360  
QY 3331 GATCAGCGGATCACCGAGTCTCCAGGTGTGGAGCTGGAGTGGAGTGAAGAGCACAGGCT 3390  
Db 3361 GATCAGCGGATCACCGAGTCTCCAGGTGTGGAGCTGGAGTGGAGTGAAGAGCACAGGCT 3420  
QY 3391 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGAGCTCAAGGCGCAGAGCCTCTCT 3450  
Db 3421 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGAGCTCAAGGCGCAGAGCCTCTCT 3480  
QY 3451 GACAGCTCAATGACTGGAGAGAGCATGTATGCTTGAATGAATGCCCGAAGCTTA 3510  
Db 3481 GACAGCTCAATGACTGGAGAGAGCATGTATGCTTGAATGAATGCCCGAAGCTTA 3540  
QY 3511 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGCCAAA 3570  
Db 3541 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGCCAAA 3600  
QY 3571 TTACAGCAGCAGATGGACCTGAGAGAAATACATTTTCCGCTCTGACTCAAGGACTGCAA 3630  
Db 3601 TTACAGCAGCAGATGGACCTGAGAGAAATACATTTTCCGCTCTGACTCAAGGACTGCAA 3660  
QY 3631 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGTGAAGTCTGAGATACAGCTG 3690  
Db 3661 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGTGAAGTCTGAGATACAGCTG 3720  
QY 3691 GAAAAATTCAGGTTCTCTATTTCTCATGAAAAAGTGAATGGAAGGCACTATTCTCAA 3750  
Db 3721 GAAAAATTCAGGTTCTCTATTTCTCATGAAAAAGTGAATGGAAGGCACTATTCTCAA 3780  
QY 3751 CAAACCAAACTCATTTGATTTCTGACGCAAAATGGAACCACTGCTTAAAGAAAAAG 3810  
Db 3781 CAAACCAAACTCATTTGATTTCTGACGCAAAATGGAACCACTGCTTAAAGAAAAAG - - 3838  
QY 3811 GGTTTATTAGTCGAGGAAGAGGACCTGCTTTACCCACACAGGTTCTCTCAGTAC 3870  
Db 3839 -----AGGTTCTCTCAGTAC 3855

QY 3871 AATGAGCTGAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGCAGAGCTAGAGAGCC 3930  
Db 3856 AATGAGCTGAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGCAGAGCTAGAGAGCC 3915  
QY 3931 CTTTCAAGAGACCCGATCGAGCTCCGGTCCGCGCGGAGGAAGTGCCTCCACCGCAAGCA 3990  
Db 3916 CTTTCAAGAGACCCGATCGAGCTCCGGTCCGCGCGGAGGAAGTGCCTCCACCGCAAGCA 3975  
QY 3991 ACGGACCAACCAACCCATCCAGCCAGCACACCGCGAGGAGAGATCGCCATTCGCGC 4050  
Db 3976 ACGGACCAACCAACCCATCCAGCCAGCACACCGCGAGGAGAGATCGCCATTCGCGC 4035  
QY 4051 ATCGTGGGCTCGCGAGAGCACCAGCCAGTGCATGAGCTGTGCGCCCGCCCATTCAGC 4110  
Db 4036 ATCGTGGGCTCGCGAGAGCACCAGCCAGTGCATGAGCTGTGCGCCCGCCCATTCAGC 4095  
QY 4111 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTTAAGGAACGATGCAC 4170  
Db 4096 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTTAAGGAACGATGCAC 4155  
QY 4171 CACAATATTTCTCACCGATTCAAGTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4230  
Db 4156 CACAATATTTCTCACCGATTCAAGTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4215  
QY 4231 TGTCTGATACCGTGCACTTTTGGACGCCAGGCATCCAAATGTCTCGAATCTCAGGTGATG 4290  
Db 4216 TGTCTGATACCGTGCACTTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCAAGTGTG 4275  
QY 4291 TGTCAACCCCAAGTCTCCAGTGTCTTCCAGCCACCTGCGGCTTGCCTGTGATATGCC 4350  
Db 4276 TGTCAACCCCAAGTCTCCAGTGTCTTCCAGCCACCTGCGGCTTGCCTGTGATATGCC 4335  
QY 4351 ACACATTTCCACCGAGGCTTCTGCGGTGACAAATGAATCCACAGGTCTCCAGACCAAG 4410  
Db 4336 ACACATTTCCACCGAGGCTTCTGCGGTGACAAATGAATCCACAGGTCTCCAGACCAAG 4395  
QY 4411 GAGCCAGCAGCAGCTTGCACCTGGAAGGTGATGAAGTGCCTCCAGGAATAACAAACGA 4470  
Db 4396 GAGCCAGCAGCAGCTTGCACCTGGAAGGTGATGAAGTGCCTCCAGGAATAACAAACGA 4455  
QY 4471 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTGAGGGATCAAAAGTCTCATTTAT 4530  
Db 4456 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTGAGGGATCAAAAGTCTCATTTAT 4515  
QY 4531 GACAAATGAACCCAGAGAGCTGACAGAGCCCGGTGGAAGAATTTGAGTGTGCTTCCC 4590  
Db 4516 GACAAATGAACCCAGAGAGCTGACAGAGCCCGGTGGAAGAATTTGAGTGTGCTTCCC 4575  
QY 4591 GACGGGATGTATCTATTCTATGCTGCGTGTGCTTCCGAATCTCGAATAACAGCCAAA 4650  
Db 4576 GACGGGATGTATCTATTCTATGCTGCGTGTGCTTCCGAATCTCGAATAACAGCCAAA 4635  
QY 4651 GCA----- 4653  
Db 4636 GCAGATGCCCATACATACCTGAAGATGAATCTCACCCGACACACCTGCTGCGCCGGG 4695  
QY 4654 ----- 4653  
Db 4696 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAAACAGCGCTGGGTCAACGCTTA 4755  
QY 4654 -----GAAAAAGCAGAAAGTCTTAGGAAAAAGCAGAGCTGATGCTAAACTG 4680  
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAAACTG 4815  
QY 4681 CTTGGAAAATCTCCCTGCTGAAAATCTGGAAGGTGATGACCGTCTAGACATGAATGCACGCTG 4740  
Db 4816 CTTGGAAAATCTCCCTGCTGAAAATCTGGAAGGTGATGACCGTCTAGACATGAATGCACGCTG 4875  
QY 4741 CCCTTCAGTGACAGGTGTGTTGTGGGACCGAGGAAGGCTCTACGCCCTGATATGC 4800  
Db 4876 CCCTTCAGTGACAGGTGTGTTGTGGGACCGAGGAAGGCTCTACGCCCTGATATGC 4935  
QY 4801 TTGAAAAATCTCCCTAAACCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860

Db 4936 TTGAAAATCTCCCTAACCCATGTCCAGGAATGGAGCAGTCTTCCAAATTTATATATC 4995  
QY 4861 AAGGACCTGGAGAAGTACTCATGATAGCAGAGAAGAGCGGGCACTGTGTCTTGTGAC 4920  
Db 4996 AAGGACCTGGAGAAGTACTCATGATAGCAGAGAAGAGCGGGCACTGTGTCTTGTGAC 5055  
QY 4921 GTGAGAAAAGTGAACAGTCCCTGGCCAGTCCCACTCCCTGGCCAGCCGACATCTCA 4980  
Db 5056 GTGAGAAAAGTGAACAGTCCCTGGCCAGTCCCACTCCCTGGCCAGCCGACATCTCA 5115  
QY 4981 CCCAAATTTTGAAGCTGTCAAGGCTGCCACTTTGTTGGGCGAGGCAAGATTGAGAAC 5040  
Db 5116 CCCAAATTTTGAAGCTGTCAAGGCTGCCACTTTGTTGGGCGAGGCAAGATTGAGAAC 5175  
QY 5041 GGGCTCTGCATCTGTGAGCCCATGCCAGCAAGTGTCTTCTCCGCTTACAAAGAAAAC 5100  
Db 5176 GGGCTCTGCATCTGTGAGCCCATGCCAGCAAGTGTCTTCTCCGCTTACAAAGAAAAC 5235  
QY 5101 CTCAGCAATATCTGATCCGGAAGAGATAGACCTCAGAGCCCTCAGCTGTATCCAC 5160  
Db 5236 CTCAGCAATATCTGATCCGGAAGAGATAGACCTCAGAGCCCTCAGCTGTATCCAC 5295  
QY 5161 TTCACCAATATACAGTATCTCTTGAACCAATTAATTTCTAGAAATCGACATGAAGCAG 5220  
Db 5296 TTCACCAATATACAGTATCTCTTGAACCAATTAATTTCTAGAAATCGACATGAAGCAG 5355  
QY 5221 TAGACGCTCAGGAATCTCGATAGATAGATGACCATCTCTGGACCTCTGTGTGCCC 5280  
Db 5356 TAGACGCTCAGGAATCTCGATAGATAGATGACCATCTCTGGACCTCTGTGTGCCC 5415  
QY 5281 GCCTCTTCAACAGCTTCCCTGTCTCAATCGTGCAGTGAACAGCGCAGGCGAGGAG 5340  
Db 5416 GCCTCTTCAACAGCTTCCCTGTCTCAATCGTGCAGTGAACAGCGCAGGCGAGGAG 5475  
QY 5341 GAGTACTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTTACGGAAGAGCTAGC 5400  
Db 5476 GAGTACTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTTACGGAAGAGCTAGC 5535  
QY 5401 CGCAGACGATCTCAAGTGGAGTCCCTTACCTTGGCCCTTGGCTTACAGAGAACCCTAT 5460  
Db 5536 CGCAGACGATCTCAAGTGGAGTCCCTTACCTTGGCCCTTGGCTTACAGAGAACCCTAT 5595  
QY 5461 CTGTTTGTGACCCACTCAACTCACTCGAAGTAATTGAGATCCAGCACGCTCCTCAGCA 5520  
Db 5596 CTGTTTGTGACCCACTCAACTCACTCGAAGTAATTGAGATCCAGCACGCTCCTCAGCA 5655  
QY 5521 GGGACCCCTCCGAGGTTACCTGGACATCCGAAACCCGGCTACCTGGGCCCTGCCATT 5580  
Db 5656 GGGACCCCTCCGAGGTTACCTGGACATCCGAAACCCGGCTACCTGGGCCCTGCCATT 5715  
QY 5581 TCCTCAGGACGATTTACTTGGGCTCTCATACAGGATAAATTAAGGTTCATTTGCTGC 5640  
Db 5716 TCCTCAGGACGATTTACTTGGGCTCTCATACAGGATAAATTAAGGTTCATTTGCTGC 5775  
QY 5641 AAGGGAACTCTGTAGGAGTCCGCACTGAACACACGGGGCCCTGCACCTCCCGC 5700  
Db 5776 AAGGGAACTCTGTAGGAGTCCGCACTGAACACACGGGGCCCTGCACCTCCCGC 5835  
QY 5701 AGCAGCCCAACAGGAGGAGGAGCCACACCTGATCAACAGGACATCAACAGCGGTGGCC 5760  
Db 5836 AGCAGCCCAACAGGAGGAGGAGCCACACCTGATCAACAGGACATCAACAGCGGTGGCC 5895  
QY 5761 TCCAGCCCAAGCGCCGAGGAGGAGCCACACCTGATCAACAGGAGCAACCCACCGC 5820  
Db 5896 TCCAGCCCAAGCGCCGAGGAGGAGCCACACCTGATCAACAGGAGCAACCCACCGC 5955  
QY 5821 TACCGCGGGGCGGACCGAGTCCGAGGAGCAAGTCTCTTGGCCGCCCTCGAGGCA 5880  
Db 5956 TACCGCGGGGCGGACCGAGTCCGAGGAGCAAGTCTCTTGGCCGCCCTCGAGGCA 6015  
QY 5881 GAGAAGTCCCGCCGCGATGCTCAGCACGCGGAGAGCGTCCCGCGGAGGCTGT 5940

Db 6016 GAGAGTCCCCCGCGGATACTCAGACCGCGAGAGCGGTCCCCCGAGGCTGTTT 6075  
QY 5941 GAAGACAGCAGAGCGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGGTG 6000  
Db 6076 GAAGACAGCAGAGCGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGGTG 6135  
QY 6001 AACAGGGAGAGGGCAGAGTGC 6023  
Db 6136 AACAGGTCTGGACCAAGTCTTC 6158  
RESULT 7  
US-10-262-511-1  
; Sequence 1, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Beghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: Curasequest version 0.1  
; SEQ ID NO 1

```
; LENGTH: 6189
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6159)
US-10-262-511-1

Query Match      85.7%; Score 5631; DB 13; Length 6189;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;

QY 19 ATGTTGAAGTCAAAATATGAGCGGGGAATCCTTTGGATGCTGGTCTGCTGAACCCCAAT 78
Db 1 ATGTTGAAGTCAAAATATGAGCGGGGAATCCTTTGGATGCTGGTCTGCTGAACCCCAAT 60
QY 79 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATGACTCAA 138
Db 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATGACTCAA 120
QY 139 CAGCAGATGTCCTCTCTTCCGAGAAAGGATATAGATGCCCTCTTTGTTCTCTTTGAA 198
Db 121 CAGCAGATGTCCTCTCTTCCGAGAAAGGATATAGATGCCCTCTTTGTTCTCTTTGAA 180
QY 199 GAATGAGTCAAGCTGCTCTGATGAAGATTAAGCAGTGAACAATTTGTCGGGAAGTAT 258
Db 181 GAATGAGTCAAGCTGCTCTGATGAAGATTAAGCAGTGAACAATTTGTCGGGAAGTAT 240
QY 259 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCCTTCGCAAGAGCTTCGAACTGAGA 318
Db 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCCTTCGCAAGAGCTTCGAACTGAGA 300
QY 319 AGTCTGTAGGTGTGGTCACTTTCTGCAAGTGCAGGTGGTGAAGAGAGAAACACCGGG 378
Db 301 AGTCTGTAGGTGTGGTCACTTTCTGCAAGTGCAGGTGGTGAAGAGAGAAACACCGGG 360
QY 379 GACATCTATGCTATGAAGTGAATGAAGAAGAGCTTTATGGCCCGAGGACAGTTTCA 438
Db 361 GACATCTATGCTATGAAGTGAATGAAGAAGAGCTTTATGGCCCGAGGACAGTTTCA 420
QY 439 TTTTGTGAGGAGAGCGGAACATATATCTCGAAGCACAAGCCGTGGATCCCCCAATTA 498
Db 421 TTTTGTGAGGAGAGCGGNAATATATCTCGAAGCACAAGCCGTGGATCCCCCAATTA 480
QY 499 CAGTATGCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGATATCAGCCTGGAGGG 558
Db 481 CAGTATGCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGATATCAGCCTGGAGGG 540
QY 559 GACTTGTCTGCTACTTTTGAATAGATGAGGACAGTTAGATGAACCTGATACAGTTT 618
Db 541 GACTTGTCTGCTACTTTTGAATAGATGAGGACAGTTAGATGAACCTGATACAGTTT 600
QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTCCACAGCTTCATCTGATGGGATACGTGCATCGA 678
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTCCACAGCTTCATCTGATGGGATACGTGCATCGG 660
QY 679 GACATCAAGCTTGAGAAATCTCTGTTGACCGCAGGACACATCAAGCTGGTGAATTT 738
Db 661 GACATCAAGCTTGAGAAATCTCTGTTGACCGCAGGACACATCAAGCTGGTGAATTT 720
QY 739 GGATCTGCCCGGAAAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTGGGACC 798
Db 721 GGATCTGCCCGGAAAATGAATTCAAACAA--GGTGAATGCCAACTCCCGATTGGGACC 777
QY 799 CCAGATTACATGGCTCTGAGAGTGTGATGATGAACCGGGATGGAAGGACACCTTAC 858
Db 778 CCAGATTACATGGCTCTGAGAGTGTGATGATGAACCGGGATGGAAGGACACCTTAC 837
QY 859 GSCCTGGAGTGTGACTGGTGTGCTGAGGCGGTGATGCTGATGAGATGATTTATGGGAGA 918
Db 838 GSCCTGGAGTGTGACTGGTGTGCTGAGGCGGTGATGCTGATGAGATGATTTATGGGAGA 897
QY 919 TCCCCCTTCGAGAGGGAACTCTGCCAGAACCTTCAATAACATTATGAATTCACGCGG 978
```

```
Db 898 TCCCCCTTCGAGAGGGAACTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 957
QY 979 TTTTGAATTTCCAGATGACCCCAAAAGTGAGCAGTGACTTTCTTGATCTGATCAAAGC 1038
Db 958 TTTTGAATTTCCAGATGACCCCAAAAGTGAGCAGTGACTTTCTTGATCTGATCAAAGC 1017
QY 1039 TTGTTGTGCGGCAGAAAGAGAGACTGAAGTTGAAGGTCTTTGCTGCCATCTTTCTTC 1098
Db 1018 TTGTTGTGCGGCAGAAAGAGAGACTGAAGTTGAAGGTCTTTGCTGCCATCTTTCTTC 1077
QY 1099 TCTAAAATTTGACTGGAAACAACATTCGTAACTCTCCTCCCCCTTCGTTCCACACCTCAAG 1158
Db 1078 TCTAAAATTTGACTGGAAACAACATTCGTAACTCTCCTCCCCCTTCGTTCCACACCTCAAG 1137
QY 1159 TCTGACGATGACACCTTCCAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTCACTCT 1218
Db 1138 TCTGACGATGACACCTTCCAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTCACTCT 1197
QY 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGACTGCGCTTTGTTGGGGTTTTCG 1278
Db 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGACTGCGCTTTGTTGGGGTTTTCG 1257
QY 1279 TACAGCAAGGCACTGGGATTTCTTTGTAGATCTGAGTCTGTTGTGCGGTCTGGACTCC 1338
Db 1258 TACAGCAAGGCACTGGGATTTCTTTGTAGATCTGAGTCTGTTGTGCGGTCTGGACTCC 1317
QY 1339 CCTGCCAAGACTAGCTCCATGGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1398
Db 1318 CCTGCCAAGACTAGCTCCATGGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1377
QY 1399 TCTCAGACAAAGTGTCAAAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1458
Db 1378 TCTCAGACAAAGTGTCAAAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1437
QY 1459 GAGGTGAGGCTGTCTTAGTCAAGAGGAGTGAGCTGAAGCCCTCTGAGACTCAGAGA 1518
Db 1438 GAGGTGAGGCTGTCTTAGTCAAGAGGAGTGAGCTGAAGCCCTCTGAGACTCAGAGA 1497
QY 1519 TCCCTCTCGAGCAGGACCTTGTCTACTCATCACAGAATGCAGTAGCTTAAAGCCGAAGT 1578
Db 1498 TCCCTCTCGAGCAGGACCTTGTCTACTCATCACAGAATGCAGTAGCTTAAAGCCGAAGT 1557
QY 1579 TTGGAGCAAGCAAGATGGAGTGTCCAGGAGGATGAACAAACACTGACAGCTTCTCCAT 1638
Db 1558 TTGGAGCAAGCAAGATGGAGTGTCCAGGAGGATGAACAAACACTGACAGCTTCTCCAT 1617
QY 1639 GATATCAGAGCAGAGCCGGRAGCTCCAGAAATCAABAGCAGGAGTACCAAGGCTCAA 1698
Db 1618 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAABAGCAGGAGTACCAAGGCTCAA 1677
QY 1699 GTGAAAGAAATGAGTTGATGATGAATCAAGTTTGAAGAGAGATCTTGTCTCAGCAAGA 1758
Db 1678 GTGAAAGAAATGAGTTGATGATGAATCAAGTTTGAAGAGAGATCTTGTCTCAGCAAGA 1737
QY 1759 CGAGTGTATCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1818
Db 1738 CGAGTGTATCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1797
QY 1819 CGAAAGCAGACAAATGTGAGCAATAAATCTTGAAGCTTAAAGATCAAGGAGAGCTGAA 1878
Db 1798 CGAAAGCAGACAAATGTGAGCAATAAATCTTGAAGCTTAAAGATCAAGGAGAGCTGAA 1857
QY 1879 GTGGGAAATATCGGAAATCTGGAGAAAGATCAATGTCTGAGCAGAGCTCAAAATTCAGGAG 1938
Db 1858 GTGGGAAATATCGGAAATCTGGAGAAAGATCAATGTCTGAGCAGAGCTCAAAATTCAGGAG 1917
QY 1939 CTCCAAGAGAACTGGAGAGGCTG----- 1963
Db 1918 CTCCAAGAGAACTGGAGAGGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGTGCGAG 1977
QY 1964 -----CAAGAGCGAGCGGAGGAGCTGGAGAGCTGGAGAGCTGAGAACCGGAGAG 2010
```

Db 1978 AATATCCGCCAAGAGCGAGCGGAGAGGAGCTGGAGAACTGCAGAACCCGAGAG 2037  
QY 2011 GATTCCTTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGGAACCGCGCCATTCTCTG 2070  
Db 2038 GATTCCTTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGGAACCGCGCCATTCTCTG 2097  
QY 2071 GAGAAACAAGGTAAAGAGACTAGAGACCATGGAGCGTGTAGAGAAACAGACTGAAGATGAC 2130  
Db 2098 GAGAAACAAGGTAAAGAGACTAGAGACCATGGAGCGTGTAGAGAAACAGACTGAAGATGAC 2157  
QY 2131 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2190  
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2217  
QY 2191 AAACATCGGAGGCCCAAGTCTCAGCCCGACCACTTGAAGTGCACCTGNAACAGAAAGAG 2250  
Db 2218 AAACATCGGAGGCCCAAGTCTCAGCCCGACCACTTGAAGTGCACCTGNAACAGAAAGAG 2277  
QY 2251 CAGCACTATCAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2310  
Db 2278 CAGCACTATCAGGAAAGATTAAGTGTGGCAATCAGATTAAGAAAGACCTGGCTGAC 2337  
QY 2311 AAGGAGACACTGAGAAACATGATGACAGACACGAGGAGGAGGCCATGAGAAAGGCAAA 2370  
Db 2338 AAGGAGACACTGAGAAACATGATGACAGACACGAGGAGGAGGCCATGAGAAAGGCAAA 2397  
QY 2371 ATTCTCAGCGAAACAGAAAGCGATGATCAATGCTATGGATTCCAAAGATCAGATCCCTGGAA 2430  
Db 2398 ATTCTCAGCGAAACAGAAAGCGATGATCAATGCTATGGATTCCAAAGATCAGATCCCTGGAA 2457  
QY 2431 CAGAGATTGTGAACTGTCTGAAGCCATAAATTTGACGCAAAATAGCAGTCTTTTACC 2490  
Db 2458 CAGAGATTGTGAACTGTCTGAAGCCATAAATTTGACGCAAAATAGCAGTCTTTTACC 2517  
QY 2491 CAAGGAACTAGAGGCCCAAGAGAGATGATTCTGAACTCAGGCAACAGAAATTTTAC 2550  
Db 2518 CAAGGAACTAGAGGCCCAAGAGAGATGATTCTGAACTCAGGCAACAGAAATTTTAC 2577  
QY 2551 CTGGAGACACAGCTGGGAAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAG 2610  
Db 2578 CTGGAGACACAGCTGGGAAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAG 2637  
QY 2611 AAGATCAGCCACCAAGACACAGTGCAGAGATCGCTCTGGAACCTGGAGACAGATTG 2670  
Db 2638 AAGATCAGCCACCAAGACACAGTGCAGAGATCGCTCTGGAACCTGGAGACAGATTG 2697  
QY 2671 CGGAGGTGAGTCTAGACACAGAGGAGCAAACTGGAGCTCAAGCGCCAGCTCAGAG 2730  
Db 2698 CGGAGGTGAGTCTAGACACAGAGGAGCAAACTGGAGCTCAAGCGCCAGCTCAGAG 2757  
QY 2731 CTAAGCTCTCCCTGAGGAGCGAGTGCAGTGTGACAGCCCTGAGGCTGCAGCGCG 2790  
Db 2758 CTAAGCTCTCCCTGAGGAGCGAGTGCAGTGTGACAGCCCTGAGGCTGCAGCGCG 2817  
QY 2791 GCCTCGAGAGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGAGCT 2850  
Db 2818 GCCTCGAGAGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGAGCT 2877  
QY 2851 GAAGAGGAGATCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2910  
Db 2878 GAAGAGGAGATCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2937  
QY 2911 CTTCTGTAAGCTGTACTGTATCAGACCTGGAGGCGAGCTAAACAGCTGACCCGAG 2970  
Db 2938 CTTCTGTAAGCTGTACTGTATCAGACCTGGAGGCGAGCTAAACAGCTGACCCGAG 2997  
QY 2971 GACAAGCTGAACTCAACAAACAAAATCTTCTACTTGTCCAAAACAACTCGATGAGCTTCT 3030  
Db 2998 GACAAGCTGAACTCAACAAACAAAATCTTCTACTTGTCCAAAACAACTCGATGAGCTTCT 3057  
QY 3031 GCGGCCAAGCAGAGATTGTACAACTGCGAAGTGAAGTGCAGCATCTCCGCCGGAGATC 3090  
Db 3058 GCGGCCAAGCAGAGATTGTACAACTGCGAAGTGAAGTGCAGCATCTCCGCCGGAGATC 3117

QY 3091 ACGGAACGAGATGAGCTTACAGCCAGAACCAACGATGAGGCTCTGAAGACCAAG 3150  
Db 3118 ACGGAACGAGATGAGCTTACAGCCAGAACCAACGATGAGGCTCTGAAGACCAAG 3177  
QY 3151 TGCAACCATGTGGAGGAACAGGTCAATGGATTTTGGAGCCCTAAACGATGAGCTGTAGAA 3210  
Db 3178 TGCAACCATGTGGAGGAACAGGTCAATGGATTTTGGAGCCCTAAACGATGAGCTGTAGAA 3237  
QY 3211 AAAGAGCGCAGTGGAGGCTCGAGGAGCTCTGGGTGATGAGAAATCCCAAGTTTGA 3270  
Db 3238 AAAGAGCGCAGTGGAGGCTCGAGGAGCTCTGGGTGATGAGAAATCCCAAGTTTGA 3297  
QY 3271 TGTCCGTTTCAGAGCTGACAGAAATGCTGACACCGAGAAACAGAGCAGGCGAGAGCC 3330  
Db 3298 TGTCCGTTTCAGAGCTGACAGAGATGCTGACACCGAGAAACAGAGCAGGCGAGAGCC 3357  
QY 3331 GATCAGCGGATCACCGAGTCTCCGAGGTGGTGGAGCTGGCAGTGAAGGAGCAAGGCT 3390  
Db 3358 GATCAGCGGATCACCGAGTCTCCGAGGTGGTGGAGCTGGCAGTGAAGGAGCAAGGCT 3417  
QY 3391 GAGATTCTCGCTCTGACAGAGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCTCTCT 3450  
Db 3418 GAGATTCTCGCTCTGACAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCTCTCT 3477  
QY 3451 GACAAGCTCAATGACCTGGAGAGAGAGCATGTATGCTTGAATGAATGCCCGAGCTTA 3510  
Db 3478 GACAAGCTCAATGACCTGGAGAGAGAGCATGTATGCTTGAATGAATGCCCGAGCTTA 3537  
QY 3511 CAGCAGAGCTGAGACTGAACGAGGCTCAACAGAGGCTCTGGAAGAGCAAGCAAA 3570  
Db 3538 CAGCAGAGCTGAGACTGAACGAGGCTCAACAGAGGCTCTGGAAGAGCAAGCAAA 3597  
QY 3571 TTACAGCAGCAGATGGACCTGCAAGAAAATCAATTTCCGTCTGACTCAAGGACTGCAA 3630  
Db 3598 TTACAGCAGCAGATGGACCTGCAAGAAAATCAATTTCCGTCTGACTCAAGGACTGCAA 3657  
QY 3631 GAAGCTTAGATCGGCTGATCTACTGAAGACAGAAAGTACTGAGATTCAGCTG 3690  
Db 3658 GAAGCTTAGATCGGCTGATCTACTGAAGACAGAAAGTACTGAGATTCAGCTG 3717  
QY 3691 GAAACCAATTCAGTCTCTATCTCATGAAAGGTGAAATGGAAGCACTATTTCTCAA 3750  
Db 3718 GAAACCAATTCAGTCTCTATCTCATGAAAGGTGAAATGGAAGCACTATTTCTCAA 3777  
QY 3751 CAAACCAAACTCATGATTTTCTGAAGCCAAAATGGAACCACTGCTAAAGAAAG 3810  
Db 3778 CAAACCAAACTCATGATTTTCTGAAGCCAAAATGGAACCACTGCTAAAGAAAGAA -- 3835  
QY 3811 GGTTTATTTAGTCGACGGAAGAGGCCCTGCTTACCCACAGAGTTCTCTGCAAGTAC 3870  
Db 3836 -----AGGTGCTCTGCAATAC 3852  
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGTCTGTGTGACAGCTTAGAGGAGCC 3930  
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGTCTGTGTGACAGCTTAGAGGAGCC 3912  
QY 3931 CTTTCAGAGACCCGCACTGAGCTCCGTCGCCCGGAGGAAAGTCCCAACCGCAAGCA 3990  
Db 3913 CTTTCAGAGACCCGCACTGAGCTCCGTCGCCCGGAGGAAAGTCCCAACCGCAAGCA 3972  
QY 3991 ACGGACCAACCAACCCATCCAGCCAGCCACCGAGGAGGAGATCGCCATGTCCGCC 4050  
Db 3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGAGGAGGAGATCGCCATGTCTGCC 4032  
QY 4051 ATCGTGGTGGCAGAGCAGCCAGTGCATGAGCTGTGCTGCCCGCCCATCCAGC 4110  
Db 4033 ATCGTGGTGGCAGAGCAGCCAGTGCATGAGCTGTGCTGCCCGCCCATCCAGC 4092  
QY 4111 CGCAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTGGCGCTTTAAGGAACGATGCAC 4170  
Db 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTGGCGCTTTAAGGAACGATGCAC 4152

QY 4171 CACAAATATTCCTCACCGATTCACAGTAGAGTGAACATCGCAGCCACAAAGTGCTGTG 4230  
Db 4153 CACAAATATTCCTCACCGATTCACAGTAGAGTGAACATCGCAGCCACAAAGTGCTGTG 4212  
QY 4231 TGTCTGGATACCGTGCACTTTGGACCGCAGGATCCAAATGTCTCGAATGTCAGGTGATG 4290  
Db 4213 TGTCTGGATACCGTGCACTTTGGACCGCAGGATCCAAATGTCTAGAAATGTCAAGGTGATG 4272  
QY 4291 TGTCAACCCCAAGTGCTCCACAGTGTGTCAGCAGCAGCTCGCGCTTGCTGCTGGAATATGCC 4350  
Db 4273 TGTCAACCCCAAGTGCTCCACAGTGTGTCAGCAGCAGCTCGCGCTTGCTGCTGGAATATGCC 4332  
QY 4351 ACACATCTTACCGAGGCTTCTGCGGTGACAAATGAACTCCGCGAGTCTCCAGACCAAG 4410  
Db 4333 ACACATCTTACCGAGGCTTCTGCGGTGACAAATGAACTCCGCGAGTCTCCAGACCAAG 4392  
QY 4411 GAGCCAGCAGCAGCTTGCACTGGAAGGTCGATGAAGTGGCCAGGAATAACAAACGA 4470  
Db 4393 GAGCCAGCAGCAGCTTGCACTGGAAGGTCGATGAAGTGGCCAGGAATAACAAACGA 4452  
QY 4471 GGACAGCAAGGCTGGACAGGAAGTACATTTGCTGGAGGATCAAAAGTCTCCTATTAT 4530  
Db 4453 GGACAGCAAGGCTGGACAGGAAGTACATTTGCTGGAGGATCAAAAGTCTCCTATTAT 4512  
QY 4531 GACATGAGCCAGAGAGCTGGACAGAGCGCGTGGAGAAATTTGAGCTGTGCTTCCC 4590  
Db 4513 GACATGAGCCAGAGAGCTGGACAGAGCGCGTGGAGAAATTTGAGCTGTGCTTCCC 4572  
QY 4591 GACGGGATGTATCTATTATCATGTCGCTTGGCTTCCGAACTCGCAAAATACAGCCAAA 4650  
Db 4573 GACGGGATGTATCTATTATCATGTCGCTTGGCTTCCGAACTCGCAAAATACAGCCAAA 4632  
QY 4651 GCA----- 4653  
Db 4633 GCAGATGTCCCATACATACTGAAGTGAATCTCACCGGCACACCACTGCTGCGCCGGG 4692  
QY 4654 ----- 4653  
Db 4693 AGAAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCCCTTA 4752  
QY 4654 ----- 4653  
Db 4753 GAATCAGTTGCGAGGTGGAGAGTTTCTAGGGAAGAAAGCAGAACTGATGCTAAACCTG 4812  
QY 4681 CTTGGAATCTCCCTGCTGAACTGAAAGTGAATGACCGCTAGACATCAATGACAGCTG 4740  
Db 4813 CTTGGAATCTCCCTGCTGAACTGAAAGTGAATGACCGCTAGACATCAATGACAGCTG 4872  
QY 4741 CCTTTCAGTGACCAAGTGCTGTTGGTGGGCAACGAGGAAGGCTCTACGCCCTGAAATGTC 4800  
Db 4873 CCTTTCAGTGACCAAGTGCTGTTGGTGGGCAACGAGGAAGGCTCTACGCCCTGAAATGTC 4932  
QY 4801 TTGAAAACTCCCTAACCCATGTCAGGAATGAGAGAGTCTTCCAAATTTATATTATC 4860  
Db 4933 TTGAAAACTCCCTAACCCATGTCAGGAATGAGAGAGTCTTCCAAATTTATATTATC 4992  
QY 4861 AAGGACTCGGAAGTCTCATGATAGCAGAGAGAGCGGCACTGTGCTTGTGGAC 4920  
Db 4993 AAGGACTCGGAAGTCTCATGATAGCAGAGAGAGCGGCACTGTGCTTGTGGAC 5052  
QY 4921 GTGAAGAAAGTGAACAGTCTCCGCGCCAGTCCCACTGCTGCGCCAGCCGACATCTCA 4980  
Db 5053 GTGAAGAAAGTGAACAGTCTCCGCGCCAGTCCCACTGCTGCGCCAGCCGACATCTCA 5112  
QY 4981 CCCACATTTTTCAGCTGTCAGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5040  
Db 5113 CCCACATTTTTCAGCTGTCAGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5172  
QY 5041 GGGCTCTGATCTGTGACGCTATGCCAGCAAGTCTCATTTCTCCGCTTACAAACGAAAC 5100  
Db 5173 GGGCTCTGATCTGTGACGCTATGCCAGCAAGTCTCATTTCTCCGCTTACAAACGAAAC 5232  
QY 5101 CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTTCAGACCTCTGAGCCTCTGATCCAC 5160

Db 5233 CTACGAAATACTGTCATCCGAAAGAGATAGAGACCTCAGAGCCTGACGTGTATCCAC 5292  
QY 5161 TTCAACCAATTTACAGTATCCTCATTTGGAAACCAATAAATTTCAACAAATTCGACATGAAGCAG 5220  
Db 5293 TTCAACCAATTTACAGTATCCTCATTTGGAAACCAATAAATTTCAACAAATTCGACATGAAGCAG 5352  
QY 5221 TACAAGCTCGAGAAATTCCTGGATAAGAAATGACCAATTCCTTGGACCTGCTGTTGTTGCC 5280  
Db 5353 TACAAGCTCGAGAAATTCCTGGATAAGAAATGACCAATTCCTTGGACCTGCTGTTGTTGCC 5412  
QY 5281 GCCTCTTCCAAACAGCTTCCCTGCTCAATCGTCAAGTGAACAGCGCAGGCGAGCGAGAG 5340  
Db 5413 GCCTCTTCCAAACAGCTTCCCTGCTCAATCGTCAAGTGAACAGCGCAGGCGAGCGAGAG 5472  
QY 5341 GAGTACTTGTGCTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC 5400  
Db 5473 GAGTACTTGTGCTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC 5532  
QY 5401 CGACAGACGATCTCAAGTGGAGTCGTTTACCTTTGGCCTTTGGCCTTACAGAGAACCTTAT 5460  
Db 5533 CGACAGACGATCTCAAGTGGAGTCGTTTACCTTTGGCCTTTGGCCTTACAGAGAACCTTAT 5592  
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCACGCTCCTCAGCA 5520  
Db 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCACGCTCCTCAGCA 5652  
QY 5521 GGAACCCCTCCCGAGCGTACTTGGCAATCCCGAAACCCGCGCTACTTGGGCCCTGCGCAAT 5580  
Db 5653 GGAACCCCTCCCGAGCGTACTTGGCAATCCCGAAACCCGCGCTACTTGGGCCCTGCGCAAT 5712  
QY 5581 TCCTCAGGACGATTTTACTTGGCGTCTCATACAGGATAAATTAAGGTCATTTGCTGTC 5640  
Db 5713 TCCTCAGGACGATTTTACTTGGCGTCTCATACAGGATAAATTAAGGTCATTTGCTGTC 5772  
QY 5641 AAGGAAACCTCTGTAAGAGTCCGCGACTGAACACCCGCGGCGCGCTCCACTCCCGC 5700  
Db 5773 AAGGAAACCTCTGTAAGAGTCCGCGACTGAACACCCGCGGCGCGCTCCACTCCCGC 5832  
QY 5701 AGACGCCCCAAAGAGAGCGCCACCCACGATCAACAGAGCAGATCAACAGAGCGGTGGCC 5760  
Db 5833 AGCAGCCCCAAAGAGAGCGCCACCCACGATCAACAGAGCAGATCAACAGAGCGGTGGCC 5892  
QY 5761 TCAGCCAGCGCGCGCGCGAGGCGCCAGCCACCCGCGAGAGCAAGCAGACACCCACCGC 5820  
Db 5893 TCAGCCAGCGCGCGCGCGAGGCGCCAGCCACCCGCGAGAGCAAGCAGACACCCACCGC 5952  
QY 5821 TACCGAGGGGCGGACCGAGCTGCGAGGGAACAAGTCTCTGCGCGCGCGCGCTGGAGCGA 5880  
Db 5953 TACCGAGGGGCGGACCGAGCTGCGAGGGAACAAGTCTCTGCGCGCGCGCGCTGGAGCGA 6012  
QY 5881 GAGAAGTCCCGCGCGGATGCTCAGCAGCGAGAGAGCGGTCCCGCGGAGGCTGTTT 5940  
Db 6013 GAGAAGTCCCGCGCGGATGCTCAGCAGCGAGAGAGCGGTCCCGCGGAGGCTGTTT 6072  
QY 5941 GAAGACAGCAGGCGCGGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCCGAGGTG 6000  
Db 6073 GAAGACAGCAGGCGCGGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCCGAGGTG 6132  
QY 6001 AACAGGGAAGAGGCGAGGTCCTCTCAGTCTTTCACCGTTTAACTGTCACCT 6055  
Db 6133 AACAGGTCGGGACCGAGTCTTTCAGTATAATCTCAGCCAGAAAAACCAACTCTT 6187

RESULT 8  
US-09-964-956-10  
; Sequence 10, Application US/09964956  
; Publication No. US20040043926A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle

APPLICANT: Stone, David  
APPLICANT: Gunther, Erik  
APPLICANT: Ellerman, Karen  
APPLICANT: Grosse, William M  
APPLICANT: Alsbrook II, John P  
APPLICANT: Lepley, Denise M  
APPLICANT: Burgess, Catherine E  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Spyttek, Kimberly A  
APPLICANT: Leach, Martin D  
APPLICANT: Shimkets, Richard A  
TITLE OF INVENTION: No. US20040043926A1el Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-124  
CURRENT APPLICATION NUMBER: US/09/964,956  
CURRENT FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/235,631  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/235,633  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/235,808  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/236,064  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/236,065  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/236,066  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/236,135  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: 60/237,434  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/238,321  
PRIOR FILING DATE: 2000-10-05  
PRIOR APPLICATION NUMBER: 60/238,399  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/238,396  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/276,667  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/294,823  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: 60/304,868  
NUMBER OF SEQ ID NOS: 127  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 10  
LENGTH: 6189  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-964-956-10

Query Match 85.7%; Score 5631; DB 13; Length 6189;  
Best Local Similarity 95.3%; Pred. No. 0; Mismatches 45; Indels 246; Gaps 4;  
Matches 5944; Conservative 0

QY 19 ATGTGAAGTTCAATATATGAGCGGGAATCTTTGGATGCTGGTGTGCTGCTGAACCCATT 78  
Db 1 ATGTGAAGTTCAATATATGAGCGGGAATCTTTGGATGCTGGTGTGCTGCTGAACCCATT 60

QY 79 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCCCTTTATGACTCAA 138  
Db 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCCCTTTATGACTCAA 120

QY 139 CAGCAGATGTCCTCTTCTCCGAGAGGGAATATAGATGCCCTCTTGTGTTCTCTTTGAA 198  
Db 121 CAGCAGATGTCCTCTTCTCCGAGAGGGAATATAGATGCCCTCTTGTGTTCTCTTTGAA 180

QY 199 GAATGAGTCAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCCGGAAGTAT 258  
Db 181 GAATGAGTCAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCCGGAAGTAT 240

QY 259 TCCGACACCATAGCTGAGTTACAGGAGTCCAGCCTTCGGCAAGGACTTCCGAAGTCAGA 318

Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTTCGAAGTCAGA 300  
QY 319 AGTCTTGTAGTTCGTGCTCACTTTGCTGAAGTCAGGTGTTAAGAGAGAAAGCAACCGGG 378  
Db 301 AGTCTTGTAGTTCGTGCTCACTTTGCTGAAGTCAGGTGTTAAGAGAGAAAGCAACCGGG 360  
QY 379 GACATCTATGCTATGAAGTGAATGAAGAAGGCTTTATTGGCCCCAGAGAGAGTTTCA 438  
Db 361 GACATCTATGCTATGAAGTGAATGAAGAAGGCTTTATTGGCCCCAGAGAGAGTTTCA 420  
QY 439 TTTTTCGAGAGAGCGGACATATATCTCGAGACACAGCCCGTGGATCCCCCAATTA 498  
Db 421 TTTTTCGAGAGAGCGGACATATATCTCGAGACACAGCCCGTGGATCCCCCAATTA 480  
QY 499 CAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 558  
Db 481 CAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGCTGATGGAATATCAGCCTGGAGGG 540  
QY 559 GACTTGTCTGCTATTTTGAATAGATATAGAGACAGTTAGATGAACACCTGATACAGTTT 618  
Db 541 GACTTGTCTGCTATTTTGAATAGATATAGAGACAGTTAGATGAACACCTGATACAGTTT 600  
QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGGCTTTCATCTGATGGGATACGTCGATCGA 678  
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGGCTTTCATCTGATGGGATACGTCGATCGG 660  
QY 679 GACATCAAGCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTGGATTTT 738  
Db 661 GACATCAAGCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTGGATTTT 720  
QY 739 GGATCTCCCGGAAATGAATTCACAAACAGATCGTGAATCCCAAACCTCCCGATTGGGACC 798  
Db 721 GGATCTCCCGGAAATGAATTCACAAACAGATCGTGAATTCACAAACCTCCCGATTGGGACC 777  
QY 799 CCAGATTACATGGCTCCTGAAAGTGTGACTGTGATGAACGGGATGAAAAGGACACTAC 858  
Db 778 CCAGATTACATGGCTCCTGAAAGTGTGACTGTGATGAACGGGATGAAAAGGACACTAC 837  
QY 859 GGCTTGACTGTGACTGTGCTGAGTGGGCTGATTCCTATGAGATGATTTATGGAGA 918  
Db 838 GGCTTGACTGTGACTGTGCTGAGTGGGCTGATTCCTATGAGATGATTTATGGAGA 897  
QY 919 TCCCCCTTCGACAGAGGAACCTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 978  
Db 898 TCCCCCTTCGACAGAGGAACCTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 957  
QY 979 TTTTGTGAATTTCCAGATGACCCCAAGTAGCAGTGAATTTCTTCTGATCTGATTCGAAGC 1038  
Db 958 TTTTGTGAATTTCCAGATGACCCCAAGTAGCAGTGAATTTCTTCTGATCTGATTCGAAGC 1017  
QY 1039 TTGTTGCGGCCAGAGAGAGACTGAAGTTTGAAGGCTTTTGTCTGCCATCTCTTTCTTC 1098  
Db 1018 TTGTTGCGGCCAGAGAGAGACTGAAGTTTGAAGGCTTTTGTCTGCCATCTCTTTCTTC 1077  
QY 1099 TCTAAAATTGACTGGAAACAACATTCGTACGCTCTCCGCCCTTCGTTCCACCCCTCAAG 1158  
Db 1078 TCTAAAATTGACTGGAAACAACATTCGTACGCTCTCCGCCCTTCGTTCCACCCCTCAAG 1137  
QY 1159 TCTGACGATGACACCTTCAATTTTGTGATGAACAGAGAGAAATTCGTGGGTTTTCATCTCT 1218  
Db 1138 TCTGACGATGACACCTTCAATTTTGTGATGAACAGAGAGAAATTCGTGGGTTTTCATCTCT 1197  
QY 1219 CCGTGCCAGCTGAGCCCTTCAGGCTTTCTCGGTTGAAGAACTGCCGTTTGTGGGTTTTCG 1278  
Db 1198 CCGTGCCAGCTGAGCCCTTCAGGCTTTCTCGGTTGAAGAACTGCCGTTTGTGGGTTTTCG 1257  
QY 1279 TACAGCAGGACCTGGGATTTCTTGTAGATCTGAGTCTGTTGTGCGGCTCTGGACTCC 1338  
Db 1258 TACAGCAGGACCTGGGATTTCTTGTAGATCTGAGTCTGTTGTGCGGCTCTGGACTCC 1317  
QY 1339 CTTGCCAAGACTAGCTCCATGGAAAAAGAACTTCTCATCAAAAGCAAGAGCTTCAAGAC 1398









QY 2731 CTACAGCTCTCCCTGACGAGCGCAGTCAAGTTGACAGCCCTCGAGCTCAACGGCG 2790  
DB |||||  
DB 2758 CTACAGCTCTCCCTGACGAGCGCAGTCAAGTTGACAGCCCTGACGCTCAACGGCG 2817  
QY |||||  
DB 2791 GCCCTGAGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAAGAGACACACAGAGCT 2850  
DB |||||  
DB 2818 GCCCTGAGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAAGAGACACACAGAGCT 2877  
QY |||||  
DB 2851 GAAGAGGAGTCCAGCACTCAACGACATAGAGTGAATCAAGCGCAAAATTTGATGCT 2910  
DB |||||  
DB 2878 GAAGAGGAGTCCAGCACTCAACGACATAGAGTGAATCAACGCGCAAAATTTGATGCT 2937  
QY |||||  
DB 2911 CTTCTGTAACAGCTGTACTGTAATCAGACAGCTGGAGGAGCTAAACAGCTGACCGAG 2970  
DB |||||  
DB 2938 CTTCTGTAACAGCTGTACTGTAATCAGACAGCTGGAGGAGCTAAACAGCTGACCGAG 2997  
QY |||||  
DB 2971 GACACGCTGAACTCAACACCAAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3030  
DB |||||  
DB 2998 GACACGCTGAACTCAACACCAAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3057  
QY |||||  
DB 3031 GGCGCCAAACGACGAGATTGTAACCTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3090  
DB |||||  
DB 3058 GGCGCCAAACGACGAGATTGTAACCTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3117  
QY |||||  
DB 3091 ACGERACGAGAGTGCAGCTTACCGCCAGAGCAAAACGATGGAGCTCTGAGACCAACG 3150  
DB |||||  
DB 3118 ACGERACGAGAGTGCAGCTTACCGCCAGAGCAAAACGATGGAGCTCTGAGACCAACG 3177  
QY |||||  
DB 3151 TGCACCATGCTGGAGAAACAGTTCATGGAATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3210  
DB |||||  
DB 3178 TGCACCATGCTGGAGAAACAGTTCATGGAATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3237  
QY |||||  
DB 3211 AAGAGCGGAGTGGAGGCGCTGGAGAGGCTCTGGGTGATGAGAAATCCAGATTGAG 3270  
DB |||||  
DB 3238 AAGAGCGGAGTGGAGGCGCTGGAGAGGCTCTGGGTGATGAGAAATCCAGATTGAG 3297  
QY |||||  
DB 3271 TGTCTGGGTTGAGAGTGCAGAGAAATGCTCGACACCGAGAAACAGAGCAGGCGAGAGCC 3330  
DB |||||  
DB 3298 TGTCTGGGTTGAGAGTGCAGAGAAATGCTCGACACCGAGAAACAGAGCAGGCGAGAGCC 3357  
QY |||||  
DB 3331 GATCAGCGATCACCCAGTCTGCCAGTGTGGAGCTGGAGCTGGAGTGAAGAGACAAAGCT 3390  
DB |||||  
DB 3358 GATCAGCGATCACCCAGTCTGCCAGTGTGGAGCTGGAGCTGGAGTGAAGAGACAAAGCT 3417  
QY |||||  
DB 3391 GAGATTCTCGCTCTGACAGAGGCTCAAAAGAGCAGAGCTGAAGCCGAGAGCTCTCT 3450  
DB |||||  
DB 3418 GAGATTCTCGCTCTGACAGAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCTCTCT 3477  
QY |||||  
DB 3451 GACAACTCAATGACCTGGAGAAAGCATGCTTATGCTTGAATGAATGCCCGAAGCTTA 3510  
DB |||||  
DB 3478 GACAACTCAATGACCTGGAGAAAGCATGCTTATGCTTGAATGAATGCCCGAAGCTTA 3537  
QY |||||  
DB 3511 CAGCAGAGCTGGAGACTGAACGAGGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAA 3570  
DB |||||  
DB 3538 CAGCAGAGCTGGAGACTGAACGAGGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAA 3597  
QY |||||  
DB 3571 TTACAGCAGCAGTGGACCTGCAGAAAAATCAATTTCCGCTCTGACTCAAGGACTGCAA 3630  
DB |||||  
DB 3598 TTACAGCAGCAGTGGACCTGCAGAAAAATCAATTTCCGCTCTGACTCAAGGACTGCAA 3657  
QY |||||  
DB 3631 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGGAGTATCAGCTG 3690  
DB |||||  
DB 3658 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGGAGTATCAGCTG 3717  
QY |||||  
DB 3691 GAAAAATTCAGGTTCTCTATTCTCATGAAAGGTGAAAAATGAAGGCACTATTCTCAA 3750  
DB |||||  
DB 3718 GAAAAATTCAGGTTCTCTATTCTCATGAAAGGTGAAAAATGAAGGCACTATTCTCAA 3777  
QY |||||  
DB 3751 CAAACCAAACTCATGATTCTTCTGCAAGCCAAATGACCAACCTGCTTAAAGAAAAAG 3810  
DB |||||  
DB 3778 CAAACCAAACTCATGATTCTTCTGCAAGCCAAATGACCAACCTGCTTAAAGAAAAA - - 3835

QY 3811 GGTTTATTTAGTCGACGGAAGAGACCCCTGCTTTTACCCACACAGGTTCTCTGACGATAC 3870  
DB |||||  
DB 3836 -----AGGTGCTCTGACGATAC 3852  
QY |||||  
DB 3871 AATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGACAGCTAGAGAAAGCC 3930  
DB |||||  
DB 3853 AATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTGACAGCTAGAGAAAGCC 3912  
QY |||||  
DB 3931 CTTTCAAGAGACCCGATCGAGCTCCGCTCGCCCGGAGGAAGTGGCCACCGCAAGCA 3990  
DB |||||  
DB 3913 CTTTCAAGAGACCCGATCGAGCTCCGCTCGCCCGGAGGAAGTGGCCACCGCAAGCA 3972  
QY |||||  
DB 3991 ACAGACACACACACACCCATCCAGCCAGCCACCGCGAGGAGAGATGCGCATCTCCGCG 4050  
DB |||||  
DB 3973 ACAGACACACACACACCCATCCAGCCAGCCACCGCGAGGAGAGATGCGCATCTCCGCG 4032  
QY |||||  
DB 4051 ATCTGTGCGTGGCCAGAGACACCGCCAGTGCCTGATGAGCTGTGGCCCGCCCATCCAGC 4110  
DB |||||  
DB 4033 ATCTGTGCGTGGCCAGAGACACCGCCAGTGCCTGATGAGCTGTGGCCCGCCCATCCAGC 4092  
QY |||||  
DB 4111 CGGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTTAAGGAACGATGAC 4170  
DB |||||  
DB 4093 CGGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTTAAGGAACGATGAC 4152  
QY |||||  
DB 4171 CACAATATTTCTCACCATTCAAGCTAGGACTGAACATGCGAGCACAAGTGTGCTGTG 4230  
DB |||||  
DB 4153 CACAATATTTCTCACCATTCAAGCTAGGACTGAACATGCGAGCACAAGTGTGCTGTG 4212  
QY |||||  
DB 4231 TGTCTGATACCGTGCACTTTGGACGCGCAGCATCCAAATGTCTGAAATGTCAGGTGATG 4290  
DB |||||  
DB 4213 TGTCTGATACCGTGCACTTTGGACGCGCAGCATCCAAATGTCTGAAATGTCAGGTGATG 4272  
QY |||||  
DB 4291 TGTCAACCCAGTGTCTCCAGTGTGCCAGCCACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 4350  
DB |||||  
DB 4273 TGTCAACCCAGTGTCTCCAGTGTGCCAGCCACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 4332  
QY |||||  
DB 4351 ACACATTTACCGAGGCTTCTGCGCTGACAAAAATGAATGAACTCCCGAGTCTCCAGACCAAG 4410  
DB |||||  
DB 4333 ACACATTTACCGAGGCTTCTGCGCTGACAAAAATGAATGAACTCCCGAGTCTCCAGACCAAG 4392  
QY |||||  
DB 4411 GAGCCAGCAGCAGCTTGCACCTGGAGAGGTGATGAAAGTGGCCCGCAGGAATAACAAAGCA 4470  
DB |||||  
DB 4393 GAGCCAGCAGCAGCTTGCACCTGGAGAGGTGATGAAAGTGGCCCGCAGGAATAACAAAGCA 4452  
QY |||||  
DB 4471 GGAAGCAAGGCTGGGACAGGAAGTCAATTTCTCTGAGGAGTCAAAAGTCTCTCATTTAT 4530  
DB |||||  
DB 4453 GGAAGCAAGGCTGGGACAGGAAGTCAATTTCTCTGAGGAGTCAAAAGTCTCTCATTTAT 4512  
QY |||||  
DB 4531 GACAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCC 4590  
DB |||||  
DB 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCC 4572  
QY |||||  
DB 4591 GACGGGATGTATCTATTCTATGCTGGCTTGGTCTTCCGAACTCGCAATACAGCCAAA 4650  
DB |||||  
DB 4573 GACGGGATGTATCTATTCTATGCTGGCTTGGTCTTCCGAACTCGCAATACAGCCAAA 4632  
QY |||||  
DB 4651 GCA ----- 4653  
DB 4633 GCAGATGTCCTATACATACATGAAATCTCACCCGACACCACTCTGCTGCGCCCGG 4692  
QY ----- 4653  
DB 4693 AGAACCTCTACTTGTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4752  
QY ----- 4653  
DB 4654 -----GAAAAAGCAGAAAGTGTGCTTAAACTG 4690  
DB 4753 GAATCAGTTGTGCGAGTGGGAGGTTTCTAGGAAAAAAGCAGAGCTGATGCTAAACTG 4812  
QY |||||  
DB 4681 CTTGGAACCTCCCTGTGTAACCTGGAAGGTGATGACGCTGTAGACATGAACCTGACCGCTG 4740  
DB |||||  
DB 4813 CTTGGAACCTCCCTGTGTAACCTGGAAGGTGATGACGCTGTAGACATGAACCTGACCGCTG 4872  
QY |||||  
DB 4741 CCCTTCAGTGACAGGTTGTTGGTGGGACCGAGGAAGGGCTCTACGCGCTGATGTC 4800

Db 4873 CCCTTCAGTGACAGGTAGTGTGGTGGGCAACGAGGAGGGCTCTACGGCCCTGAATGTC 4932  
QY TTGAAAGAACTCCCTTAACCCATGTCCTCCAGGAATGGAGCAGTCTTCCAAATTTATATATTC 4860  
Db 4933 TTGAAAGAACTCCCTTAACCCATGTCCTCCAGGAATGGAGCAGTCTTCCAAATTTATATATTC 4992  
QY AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 4920  
Db 4993 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 5052  
QY GTGAAGAAAGTGAAACAGTCCCTGGCCAGTCCCACTGGCTCCCGAGCCGACATCTCA 4980  
Db 5053 GTGAAGAAAGTGAAACAGTCCCTGGCCAGTCCCACTGGCTCCCGAGCCGACATCTCA 5112  
QY CCCAACATTTTGAAGCTGTCAAGGGGTGCCATTTGTTGGGGCAGGCAAGATTGAGAAC 5040  
Db 5113 CCNACATTTTGAAGCTGTCAAGGGGTGCCATTTGTTGGGGCAGGCAAGATTGAGAAC 5172  
QY GGGCTCTGCAATGTGTGAGCCATGCCCGAGCAAAAGTGTCTATTTCCGCTTACAACGAAAC 5100  
Db 5173 GGGCTCTGCAATGTGTGAGCCATGCCCGAGCAAAAGTGTCTATTTCCGCTTACAACGAAAC 5232  
QY CTGAGCAAAATGCTCATCCGGAAGAGATAGACCTCAGAGCCCTGCAGCTGTATCCAC 5160  
Db 5233 CTGAGCAAAATGCTCATCCGGAAGAGATAGACCTCAGAGCCCTGCAGCTGTATCCAC 5292  
QY TTCAACCAATTTACAGTATCTCTCATTTGGAACCAATAAATTTACGAAATCGACATGAAGCAG 5220  
Db 5293 TTCAACCAATTTACAGTATCTCTCATTTGGAACCAATAAATTTACGAAATCGACATGAAGCAG 5352  
QY TACAGCTCAGGAAATTCCTGGATAGATGACATTCCTTGGCACCTGCTGTTTGGC 5280  
Db 5353 TACAGCTCAGGAAATTCCTGGATAGATGACATTCCTTGGCACCTGCTGTTTGGC 5412  
QY GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAAACGAGCGGAGCGAGAG 5340  
Db 5413 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAAACGAGCGGAGCGAGAG 5472  
QY GAGTACTGTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGGATTTTACGGAAGACGTAGC 5400  
Db 5473 GAGTACTGTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGGATTTTACGGAAGACGTAGC 5532  
QY CGACAGACCATCTCAAGTGGAGTTCGCTTACCTTTGCGCTTACGCTTACAGAGAACCTAT 5460  
Db 5533 CGACAGACCATCTCAAGTGGAGTTCGCTTACCTTTGCGCTTACGCTTACAGAGAACCTAT 5592  
QY CTGTTGTGACCCACTTCAACTCACTCGAAGTAATTTAGATCCAGCACGCTCCTCAGCA 5520  
Db 5593 CTGTTGTGACCCACTTCAACTCACTCGAAGTAATTTAGATCCAGCACGCTCCTCAGCA 5652  
QY GGGACCCCTCCCGAGGTTACCTTGGCGTCTCTATACCAGGCAATTAAGGGTCAATTTGCTGC 5580  
Db 5653 GGGACCCCTCCCGAGGTTACCTTGGCGTCTCTATACCAGGCAATTAAGGGTCAATTTGCTGC 5712  
QY TCCTCAGGAGCGATTTACTTTGGCGTCTCTATACCAGGCAATTAAGGGTCAATTTGCTGC 5640  
Db 5713 TCCTCAGGAGCGATTTACTTTGGCGTCTCTATACCAGGCAATTAAGGGTCAATTTGCTGC 5772  
QY AAGGGAACCTCTGTAAGAGTTCGGGATCGGACATCGAACCCGCGGCTCCACCTCCCGC 5700  
Db 5773 AAGGGAACCTCTGTAAGAGTTCGGGATCGGACATCGAACCCGCGGCTCCACCTCCCGC 5832  
QY AGCAGCCCCAACAGGAGGCCCCACCACTCAACAGGACATCAACAGGCGGTGGCC 5760  
Db 5833 AGCAGCCCCAACAGGAGGCCCCACCACTCAACAGGACATCAACAGGCGGTGGCC 5892  
QY TCCAGCCAGCGCGCGGAGGCCCCAGCCACCCCGGAGAGCCAAAGCACCCACCGC 5820  
Db 5893 TCCAGCCAGCGCGCGGAGGCCCCAGCCACCCCGGAGAGCCAAAGCACCCACCGC 5952  
QY TACCGGAGGGGGGAGCCGAGGAGTTCGAGGAGCAAGTCTCTTGGCGGCGCCCTGGAGGCA 5880

Db 5953 TACCGGAGGGGGGAGCCGAGCTCGCAGGAGCAAGTCTCTTGGCGGCGCCCTGGAGGCA 6012  
QY 5881 GAGAAGTCCCCCGGCGGATGCTCAGCACGCGGAGAGAGCGGTCCCCCGGAGGCTGTTT 5940  
Db 6013 GAGAAGTCCCCCGGCGGATGCTCAGCACGCGGAGAGAGCGGTCCCCCGGAGGCTGTTT 6072  
QY 5941 GAAGACAGCAGCAGGGGCGGCTGCTGCTGCGGAGACCCCGCTGTCCAGGTG 6000  
Db 6073 GAAGACAGCAGCAGGGGCGGCTGCTGCTGCGGAGACCCCGCTGTCCAGGTG 6132  
QY 6001 AACAGGGAAGAGGCA 6017  
Db 6133 AACAGGTGAGGCGACGA 6149

## RESULT 10

US-09-964-956-8  
; Sequence 8, Application US/09964956  
; Publication No. US20040043926A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David  
; APPLICANT: Gunther, Erik  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Leach, Martin D  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20040043926A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-124  
; CURRENT APPLICATION NUMBER: US/09/964,956  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/235,631  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235,633  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235,808  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,064  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,065  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,066  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,135  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/237,434  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/238,321  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,399  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,396  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/276,667  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/294,823  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: 60/304,868  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 6201  
; TYPE: DNA

; ORGANISM: Homo sapiens		Query Match		85.6%; Score 5629.8; DB 13; Length 6201;	
US-09-964-956-8		Best Local Similarity		95.7%; Pred. No. 0;	
		Matches 5929; Conservative		22; Mismatches	
				22; Indels 246; Gaps 4;	
QY	19	ATGTTGAAGTTC	CAAAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAAACCCATT	78	
DB	1	ATGTTGAAGTTC	CAAAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAAACCCATT	60	
QY	79	GCAGCCGGGCTC	CAAGGTGAATCTGTTCTCCAGGGGAAACCAACCCCTTTATGACTCAA	138	
DB	61	GCAGCCGGGCTC	CAAGGTGAATCTGTTCTCCAGGGGAAACCAACCCCTTTATGACTCAA	120	
QY	139	CAGCAGATGTC	CTCTCTTCCGAGAGGATATAGATGCCCTCTTCTCTCTTTGAA	198	
DB	121	CAGCAGATGTC	CTCTCTTCCGAGAGGATATAGATGCCCTCTTCTCTCTTTGAA	180	
QY	199	GAATGAGTCTG	CTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGGAAGTAT	258	
DB	181	GAATGAGTCTG	CTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGGAAGTAT	240	
QY	259	TCCGACACCAT	AGCTGAGTACAGGAGCTCCAGCCTTCGGCAAGAGACTTCGAACTCAGA	318	
DB	241	TCCGACACCAT	AGCTGAGTACAGGAGCTCCAGCCTTCGGCAAGAGACTTCGAACTCAGA	300	
QY	319	AGTCTTGTAG	TTGTTGCTGAAGTGCAGGTGGTAAAGAGAGAAACCAACCCGG	378	
DB	301	AGTCTTGTAG	TTGTTGCTGAAGTGCAGGTGGTAAAGAGAGAAACCAACCCGG	360	
QY	379	GACATCTATG	CTATGAAGTATGAAGAAGAGGCTTTATGGCCCGAGGAGAGGTTTCA	438	
DB	361	GACATCTATG	CTATGAAGTATGAAGAAGAGGCTTTATGGCCCGAGGAGAGGTTTCA	420	
QY	439	TTTTTTGAG	GAAGAGGGAACATTTATCTCGAAGCACAAGCCGTGGATCCCCCAATTA	498	
DB	421	TTTTTTGAG	GAAGAGGGAACATTTATCTCGAAGCACAAGCCGTGGATCCCCCAATTA	480	
QY	499	CAGTATGCTT	TCAGGACAAAATACCTTTATCTGATGAGGAATATCAGCCTCGAGGG	558	
DB	481	CAGTATGCTT	TCAGGACAAAATACCTTTATCTGATGAGGAATATCAGCCTCGAGGG	540	
QY	559	GACTTGCTCT	CACTTTGATAGATAGGACCACTAGTAGTGAACCTGATCAGTTT	618	
DB	541	GACTTGCTCT	CACTTTGATAGATAGGACCACTAGTAGTGAACCTGATCAGTTT	600	
QY	619	TACCTAGCT	GAGTGAATTTGGCTGTTTACAGCGTTTCATCTGATGGATACGTGCATCGA	678	
DB	601	TACCTAGCT	GAGTGAATTTGGCTGTTTACAGCGTTTCATCTGATGGATACGTGCATCGG	660	
QY	679	GACATCAAG	CCCTGAGACATTTCTGTTGACCGCAGAGACATCAAGCTGGTGAATTT	738	
DB	661	GACATCAAG	CCCTGAGACATTTCTGTTGACCGCAGAGACATCAAGCTGGTGAATTT	720	
QY	739	GGATCTGCG	CGGAAAATGAATTCAAACAAGATGGTGAATGCCAACTCCGATTTGGGACC	798	
DB	721	GGATCTGCG	CGGAAAATGAATTCAAACAAGATGGTGAATGCCAACTCCGATTTGGGACC	777	
QY	799	CCAGATTA	CATGGCTCCTGAACTGCTGATGAACGGGATGGAAGAGGACCTTAC	858	
DB	778	CCAGATTA	CATGGCTCCTGAACTGCTGATGAACGGGATGGAAGAGGACCTTAC	837	
QY	859	GGCTCGACT	GTGACTGGTGGTTCAGTGGCGGTGATTCCTATGAGATGATTTATGGGAGA	918	
DB	838	GGCTCGACT	GTGACTGGTGGTTCAGTGGCGGTGATTCCTATGAGATGATTTATGGGAGA	897	
QY	919	TCCCCCTTC	GCAGAGGAACTCTGCGAGAACCTTCAATAACATTAATGAAATTCAGCGGG	978	
DB	898	TCCCCCTTC	GCAGAGGAACTCTGCGAGAACCTTCAATAACATTAATGAAATTCAGCGGG	957	
QY	979	TTTTTGAAT	TTTCCAGATGACCCCAAGTGACGAGTACTTTCTTGATCTGATTCGAAAGC	1038	

DB	958	TTTTGAAATTT	CCAGATGACCCCAAGTGACAGTGTCTTCTTCTGATCTGATTCGAAAGC	1017	
QY	1039	TTGTTGCGG	CCAGAAAGAGAGACTGAAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC	1098	
DB	1018	TTGTTGCGG	CCAGAAAGAGAGACTGAAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC	1077	
QY	1099	TCTAAATTT	GACTGGAAACCAATTCGTAACCTCTCTCCCTTCTGTTCCCAACCTCAAG	1158	
DB	1078	TCTAAATTT	GACTGGAAACCAATTCGTAACCTCTCTCCCTTCTGTTCCCAACCTCAAG	1137	
QY	1159	TCTGACGAT	GAACACTCTCAATTTTGTATGAACAGAGAAATTCGTGGGTTTCATCTCT	1218	
DB	1138	TCTGACGAT	GAACACTCTCAATTTTGTATGAACAGAGAAATTCGTGGGTTTCATCTCTCT	1197	
QY	1219	CCGTGCCAG	CTGAGCCCTCAGGCTTCTCGSGTCAAGAACTGCCGTTTGTGGGGTTTTTCG	1278	
DB	1198	CCGTGCCAG	CTGAGCCCTCAGGCTTCTCGSGTCAAGAACTGCCGTTTGTGGGGTTTTTCG	1257	
QY	1279	TACAGCAAG	CGACTGGGATCTTGTGTAGTCTGAGTCTGTGTGTGGGTCTGAGTCTC	1338	
DB	1258	TACAGCAAG	CGACTGGGATCTTGTGTAGTCTGAGTCTGTGTGTGGGTCTGAGTCTC	1317	
QY	1339	CCTGCCAAG	ACTAGCTCCATGGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC	1398	
DB	1318	CCTGCCAAG	ACTAGCTCCATGGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC	1377	
QY	1399	TCTCAGGAC	CAAGTGTCAAGATGGAGCAGGAAATGACCCGGTTTACATCGAGAGTGTCA	1458	
DB	1378	TCTCAGGAC	CAAGTGTCAAGATGGAGCAGGAAATGACCCGGTTTACATCGAGAGTGTCA	1437	
QY	1459	GAGTGAGG	CTGTCTAGCAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA	1518	
DB	1438	GAGTGAGG	CTGTCTAGCAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA	1497	
QY	1519	TCCCTCTCT	GTGAGCAGCTTGTCTACCTCATCATCAGAAATGCAGTACTTAAAGCGAAGT	1578	
DB	1498	TCCCTCTCT	GTGAGCAGCTTGTCTACCTCATCATCAGAAATGCAGTACTTAAAGCGAAGT	1557	
QY	1579	TTGAGCAAG	CACGATGGAGTGTCCAGAGGATGACAAAGCATCTGACGTTCTCCAT	1638	
DB	1558	TTGAGCAAG	CACGATGGAGTGTCCAGAGGATGACAAAGCATCTGACGTTCTCCAT	1617	
QY	1639	GATATCAG	AGAGAGAGCCGAGCTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAA	1698	
DB	1618	GATATCAG	AGAGAGAGCCGAGCTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAA	1677	
QY	1699	GTGGAAGAA	TGAGTTGATGAATCACTTGGAGAGGATCTTGTCTCAGCAAGAAGA	1758	
DB	1678	GTGGAAGAA	TGAGTTGATGAATCACTTGGAGAGGATCTTGTCTCAGCAAGAAGA	1737	
QY	1759	CGGAGTGA	TCTCTACGAACTGTGAGTGTGAGAGTCTCGGCTTCTGCTGAGAAATTCAG	1818	
DB	1738	CGGAGTGA	TCTCTACGAACTGTGAGTGTGAGAGTCTCGGCTTCTGCTGAGAAATTCAG	1797	
QY	1819	CGGAAACCG	CAGCAATGTGACATAAACTGTTCAAGGCTAAGGATCAAGGAAAGCCCTGAA	1878	
DB	1798	CGGAAACCG	CAGCAATGTGACATAAACTGTTCAAGGCTAAGGATCAAGGAAAGCCCTGAA	1857	
QY	1879	GTGGGAAAT	ATGCGAAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1938	
DB	1858	GTGGGAAAT	ATGCGAAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1917	
QY	1939	CTCCAGAG	AAACTGGAGAAAGGCT-----	1962	
DB	1918	CTCCAGAG	AAACTGGAGAAAGGCTGTAAAGCCAGCAGGAGCCACCGAGCTGTGCGAG	1977	
QY	1963	-----	-----GCAAGAGCGAGCCGAGAGGAGCTGGAGAGCTGAGAAACCGAGAG	2010	
DB	1978	AATATCCG	CAGGCAAGAGGAGCGAGCGAGGAGCTGGAGAACTGCGAGAAACCGAGAG	2037	
QY	2011	GATTCTTCT	GAAGGCATCAGAAAGAGCTGTGGAGCTGAGGAAACCGCCGCTTCTCTG	2070	
DB	2038	GATTCTTCT	GAAGGCATCAGAAAGAGCTGTGGAGCTGAGGAAACCGCCGCTTCTCTG	2097	

QY 2071 GAGAACAAAGGTAAGAGACTAGAGACCATGGACGCTAGAGAAACAGACTGAAGGATGAC 2130  
DB 2098 GAGAACAAAGGTAAGAGACTAGAGACCATGGACGCTAGAGAAACAGACTGAAGGATGAC 2157  
QY 2131 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAATCTGGAGCTCGAAGAG 2190  
DB 2158 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAATCTGGAGCTCGAAGAG 2217  
QY 2191 AAACATCGGAGGCCCAAGTCTCAGGCCAGCACTAGAAGTGCACCTGAAAACAGAAAGAG 2250  
DB 2218 AAACATCGGAGGCCCAAGTCTCAGGCCAGCACTAGAAGTGCACCTGAAAACAGAAAGAG 2277  
QY 2251 CAGCACTATCAGAAAGATTAAGTTGGTGGACAATCAGATAAAGAAACCTGGCTGAC 2310  
DB 2278 CAGCACTATCAGAAAGATTAAGTTGGTGGACAATCAGATAAAGAAACCTGGCTGAC 2337  
QY 2311 AAGGAGACATCGAGAACATGATCGAGAGACAGAGGAGGCCCATGAGAAGGCCAAA 2370  
DB 2338 AAGGAGACATCGAGAACATGATCGAGAGACAGAGGAGGCCCATGAGAAGGCCAAA 2397  
QY 2371 ATTCTCAGGAAACAGAAAGGCGATGATCAATGCTATGATTCGAATTCAGATCCCTGGAA 2430  
DB 2398 ATTCTCAGGAAACAGAAAGGCGATGATCAATGCTATGATTCGAATTCAGATCCCTGGAA 2457  
QY 2431 CAGAGGATTCGGAACCTGCTGAAGCCCAATAAATTCAGCAATAGCAGTCTTTTACC 2490  
DB 2458 CAGAGGATTCGGAACCTGCTGAAGCCCAATAAATTCAGCAATAGCAGTCTTTTACC 2517  
QY 2491 CAAGGAAACATGAAGGCCCAAGAGATGATTTCTGAATTCAGGCAACAGAAATTTTAC 2550  
DB 2518 CAAGGAAACATGAAGGCCCAAGAGATGATTTCTGAATTCAGGCAACAGAAATTTTAC 2577  
QY 2551 CTGAGACACAGCTGGGAATTTGGAGGCCCAAGACCGAAATCGAGGAGCAGCTGGAG 2610  
DB 2578 CTGAGACACAGCTGGGAATTTGGAGGCCCAAGACCGAAATCGAGGAGCAGCTGGAG 2637  
QY 2611 AAGATCAGGCCAACAGACAGTGAAGAACTCGGCTGCTGAATCGGAACTGGAGACAAGATTG 2670  
DB 2638 AAGATCAGGCCAACAGACAGTGAAGAACTCGGCTGCTGAATCGGAACTGGAGACAAGATTG 2697  
QY 2671 CGGAGGTCAGTCTAGAGACAGAGGAGCAAGAACTGGAGCTCAAGGCCAGCTCAGAG 2730  
DB 2698 CGGAGGTCAGTCTAGAGACAGAGGAGCAAGAACTGGAGCTCAAGGCCAGCTCAGAG 2757  
QY 2731 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTCAGAGCCCTGCAAGCTGCAAGGCG 2790  
DB 2758 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTCAGAGCCCTGCAAGCTGCAAGGCG 2817  
QY 2791 GCCTGAGAGCAGCTTCCAGGCGGAGAGACAGAGCTGGAAGAGACCAAGAGCT 2850  
DB 2818 GCCTGAGAGCAGCTTCCAGGCGGAGAGACAGAGCTGGAAGAGACCAAGAGCT 2877  
QY 2851 GAAGAGAGATCCAGGCACTCAGGACATAGAGATGAATCCAGGCAAAATTTGATGCT 2910  
DB 2878 GAAGAGAGATCCAGGCACTCAGGACATAGAGATGAATCCAGGCAAAATTTGATGCT 2937  
QY 2911 CTTCTGAACAGCTGTACTGTAAATCAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2970  
DB 2938 CTTCTGAACAGCTGTACTGTAAATCAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2997  
QY 2971 GACACGCTGAATCCACACCAAACTTACTGCTCCAAACAACTCGATGAGGCTTCT 3030  
DB 2998 GACACGCTGAATCCACACCAAACTTACTGCTCCAAACAACTCGATGAGGCTTCT 3057  
QY 3031 GGGGCCAAACAGCAGATTTGATCAACTCGCAAGTGAAGTGAACCATCTCCGCCGGGAGATC 3090  
DB 3058 GGGGCCAAACAGCAGATTTGATCAACTCGCAAGTGAAGTGAACCATCTCCGCCGGGAGATC 3117  
QY 3091 ACGGAACGAGATGAGCTTACAGCCAGAGCAACCAATGGAGGCTCTGAAGCCAG 3150  
DB 3118 ACGGAACGAGATGAGCTTACAGCCAGAGCAACCAATGGAGGCTCTGAAGCCAG 3177

QY 3151 TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA 3210  
DB 3178 TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA 3237  
QY 3211 AAAGAGCGGCAAGTGGGAGGCTCGAGGAGCGTCTCGGTGATGAGAAATCCAGTTTTCAG 3270  
DB 3238 AAAGAGCGGCAAGTGGGAGGCTCGAGGAGCGTCTCGGTGATGAGAAATCCAGTTTTCAG 3297  
QY 3271 TGTCCGGTTCCAGAGCTGCGAGAGAAATGCTGACACCGAGAAAACAGAGAGCGGCGAGAGCC 3330  
DB 3298 TGTCCGGTTCCAGAGCTGCGAGAGAAATGCTGACACCGAGAAAACAGAGAGCGGCGAGAGCC 3357  
QY 3331 GATCAGCGGATCACCGAGTCTCGCAGGTGCTGAGCTGGCAGTGGCAGTGAAGGAGCAAGGCT 3390  
DB 3358 GATCAGCGGATCACCGAGTCTCGCAGGTGCTGAGCTGGCAGTGAAGGAGCAAGGCT 3417  
QY 3391 GAGATTTCTCGTCTGCGAGCAGGCTCTCAAGAGCAGAAAGCTTGAAGGCGGAGAGCTCTCT 3450  
DB 3418 GAGATTTCTCGTCTGCGAGCAGGCTCTCAAGAGCAGAAAGCTTGAAGGCGGAGAGCTCTCT 3477  
QY 3451 GACAGCTCANTGACCTGGAGAGCAATGCTATGCTTGAATGAATGCCCCGAGGCTTAA 3510  
DB 3478 GACAGCTCANTGACCTGGAGAGCAATGCTATGCTTGAATGAATGCCCCGAGGCTTAA 3537  
QY 3511 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3570  
DB 3538 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3597  
QY 3571 TTACAGCAGAGATGAGACTGCGAGAAAATCAGATTTTCGTCTGACTCAAGGACTGCAA 3630  
DB 3598 TTACAGCAGAGATGAGACTGCGAGAAAATCAGATTTTCGTCTGACTCAAGGACTGCAA 3657  
QY 3631 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAAGTGAAGTCAAGCTG 3690  
DB 3658 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAAGTGAAGTCAAGCTG 3717  
QY 3691 GAAACATTCAGTTCTTATCTCATGAAAGGTGAAATGGAAGGCACTATTTCCTCAA 3750  
DB 3718 GAAACATTCAGTTCTTATCTCATGAAAGGTGAAATGGAAGGCACTATTTCCTCAA 3777  
QY 3751 CAACCCAACTCATTTGATTTCTGCAAGCCAAATGACCACTGCTAAAGAAAGAAAG 3810  
DB 3778 CAACCCAACTCATTTGATTTCTGCAAGCCAAATGACCACTGCTAAAGAAAGAAAG -- 3835  
QY 3811 GGTTTATTTAGTCGACGGAAGAGGACCTCTTTACCCACACAGGTTCTCTGCAATAC 3870  
DB 3836 -----AGGTGCTCTGCAATAC 3852  
QY 3871 AATGAGCTGAAGCTGCGCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGCC 3930  
DB 3853 AATGAGCTGAAGCTGCGCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGCC 3912  
QY 3931 CTTCAGAGAGACCCGATCGAGCTCCGCTCCGCCCGGAGGAAGCTGCCCAACGCAAGCA 3990  
DB 3913 CTTCAGAGAGACCCGATCGAGCTCCGCTCCGCCCGGAGGAAGCTGCCCAACGCAAGCA 3972  
QY 3991 ACGGACCAACCAACCCATCCAGCCAGCCACCGCAGGAGCAGAGATCGGCATGTCGCC 4050  
DB 3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGCAGGAGCAGAGATCGGCATGTCGCC 4032  
QY 4051 ATCGTGGCTCGCCAGAGCAGCCAGCCAGCTGAGCTGCTGGCGCCCGCCCATCCAGC 4110  
DB 4033 ATCGTGGCTCGCCAGAGCAGCCAGCCAGCTGAGCTGCTGGCGCCCGCCCATCCAGC 4092  
QY 4111 CGCAGAAAGAGTCTTCAAATTCAGAGAAATTTAGTCCGCGCTTTAAGGAAACGATGCAAC 4170  
DB 4093 CGCAGAAAGAGTCTTCAAATTCAGAGAAATTTAGTCCGCGCTTTAAGGAAACGATGCAAC 4152  
QY 4171 CACAATATTCCTCACCATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTGCTGTG 4230  
DB 4153 CACAATATTCCTCACCATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTGCTGTG 4212  
QY 4231 TGTCTGATACCGTGCATCTTTGAGCGCCAGGCAATCCAAATGCTCTGAAATGCTCAGGATG 4290

Db 4213 TGTCGGATACCGTGACCTTTGGAGCCAGGCAATCCAAATGCTAGATGTGAGGTGATG 4272  
QY 4291 TGTACCCCAAGTGTCCACGTGCTTGGCAGCCACCTGGCGTGTGCTGCTGAATATGCC 4350  
Db 4273 TGTACCCCAAGTGTCCACGTGCTTGGCAGCCACCTGGCGTGTGCTGCTGAATATGCC 4332  
QY 4351 ACACACTTCACCGAGCCCTTCTGCGTGACAAAATGAATCCCGAGGCTCCAGACCCAG 4410  
Db 4333 ACACACTTCACCGAGCCCTTCTGCGTGACAAAATGAATCCCGAGGCTCCAGACCCAG 4392  
QY 4411 GAGCCAGCAGCAGCTTGACCTTGAAGGGTGGATGAAGGTGCCAGGAAATACCAAGGA 4470  
Db 4393 GAGCCAGCAGCAGCTTGACCTTGAAGGGTGGATGAAGGTGCCAGGAAATACCAAGGA 4452  
QY 4471 GGACAGCAGGCTGGACAGGAATGATATGCTCGAGGAGATCAAAAGTCTCATTTAT 4530  
Db 4453 GGACAGCAGGCTGGACAGGAATGATATGCTCGAGGAGATCAAAAGTCTCATTTAT 4512  
QY 4531 GACAAATGAAGCAGACAGAGCTGGACAGGCGGTGGAGAAATTTGAGCTGTGCTTCCC 4590  
Db 4513 GACAAATGAAGCAGACAGAGCTGGACAGGCGGTGGAGAAATTTGAGCTGTGCTTCCC 4572  
QY 4591 GACGGGATGTATCTATTATCATGGTGCCTTGGTGTCTTCCGAACTCGCAAAATACGCCAAA 4650  
Db 4573 GACGGGATGTATCTATTATCATGGTGCCTTGGTGTCTTCCGAACTCGCAAAATACGCCAAA 4632  
QY 4651 GCA----- 4653  
Db 4633 GCAGATGTCCTACATCTGAAGATGGAATCTCACCCGCAACACCTGTGTGGCCCGGG 4692  
QY 4654 ----- 4653  
Db 4693 AGAACCTCTACTTGTAGCTCCAGCTCCCTGACAAACAGCGCTGGGTACCCGCTTA 4752  
QY 4654 ----- 4653  
Db 4753 GAATCAGTTGTCCGAGTGGAGAGTTTCTACGGGAAAGAGAGCTGATGCTAAACTG 4812  
QY 4681 CTTGGAATCTCCCTGCTGAAATCTGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4740  
Db 4813 CTTGGAATCTCCCTGCTGAAATCTGAAGGTGATGACCGTCTAGACATGAATGACGCTG 4872  
QY 4741 CCTTCAGTGACAGTGTGTGTGGGCAACGAGAAAGGCTCTACGCCCTGAATGTC 4800  
Db 4873 CCTTCAGTGACAGTGTGTGTGGGCAACGAGAAAGGCTCTACGCCCTGAATGTC 4932  
QY 4801 TTGAAATCTCCCTAACCCATGTCACAGAAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860  
Db 4933 TTGAAATCTCCCTAACCCATGTCACAGAAATTTGGAGCAGTCTTCCAAATTTATATTATC 4992  
QY 4861 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 4920  
Db 4993 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 5052  
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCGACCTGCTGCCAGCCGACATCTCA 4980  
Db 5053 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCGACCTGCTGCCAGCCGACATCTCA 5112  
QY 4981 CCCAATTTTTGAAGCTGTCAAGGCTGCCACTTGTTTGGGCGCAGCAAGATTGAGAAC 5040  
Db 5113 CCCAATTTTTGAAGCTGTCAAGGCTGCCACTTGTTTGGGCGCAGCAAGATTGAGAAC 5172  
QY 5041 GGGCTCTGATCTGTGACGCCATGCCAGCAAGTGTGATCTTCGGCTACAAAGAAAC 5100  
Db 5173 GGGCTCTGATCTGTGACGCCATGCCAGCAAGTGTGATCTTCGGCTACAAAGAAAC 5232  
QY 5101 CTCAGCAATACTGTACTCCGGAAGAGATAGAGACTCTAGAGCCCTGAGCTGTATCCAC 5160  
Db 5233 CTCAGCAATACTGTACTCCGGAAGAGATAGAGACTCTAGAGCCCTGAGCTGTATCCAC 5292  
QY 5161 TTCACCAATTTACGATCTCTTGGACCAATAAATTTCTAGCAATGCAGATGAAGCAG 5220

Db 5293 TTCACCAATTTACAGTATCTCTTATGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5352  
QY 5221 TACACGCTCGAGGAATTTCTTGGATAAGAAATGAACATTTCTTGGCACCTGTGTGTTGCC 5280  
Db 5353 TACACGCTCGAGGAATTTCTTGGATAAGAAATGAACATTTCTTGGCACCTGTGTGTTGCC 5412  
QY 5281 GCCTTTCCCAACAGCTTCCCTGTCTCAATCGTGAGGTGAACGCGCAGGCGACGGAGAG 5340  
Db 5413 GCCTTTCCCAACAGCTTCCCTGTCTCAATCGTGAGGTGAACGCGCAGGCGACGGAGAG 5472  
QY 5341 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC 5400  
Db 5473 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC 5532  
QY 5401 CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCTTTGGCTTACAGAGAACCTTAT 5460  
Db 5533 CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGGCTTTGGCTTACAGAGAACCTTAT 5592  
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCAACGCTCTCTCAGCA 5520  
Db 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCAACGCTCTCTCAGCA 5652  
QY 5521 GGGACCCCTGCGCAGCGTACCTGGACATCCCGAAACCCGCGCTACCTGGGCGCTTGCATT 5580  
Db 5653 GGGACCCCTGCGCAGCGTACCTGGACATCCCGAAACCCGCGCTACCTGGGCGCTTGCATT 5712  
QY 5581 TCCTCAGAGAGGATTTACTTTGGCGTCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5640  
Db 5713 TCCTCAGAGAGGATTTACTTTGGCGTCTCATACAGGATAAATTAAGGGTCAATTTGCTGC 5772  
QY 5641 AAGGAAACCTCTGTAAAGGATCGGCACTGAACACCAACCGGGGCGCTTCAACCTCCGC 5700  
Db 5773 AAGGAAACCTCTGTAAAGGATCGGCACTGAACACCAACCGGGGCGCTTCAACCTCCGC 5832  
QY 5701 AGCAGCCCAACAGCAGGCGCCACCCACCTGACAGACATCAACAGACATCAACAGCGCTGCC 5760  
Db 5833 AGCAGCCCAACAGCAGGCGCCACCCACCTGACAGACATCAACAGACATCAACAGCGCTGCC 5892  
QY 5761 TCCAGCCCAAGCGCCCGCGAAGCGCCAGCAACCGCGAGAGCGCAACACCCACCGC 5820  
Db 5893 TCCAGCCCAAGCGCGCCCGCGAAGCGCCAGCAACCGCGAGAGCGCAACACCCACCGC 5952  
QY 5821 TACCGCAGGGCGGACCGAGCTGCGCAGGGAAGAATCTCTGCGCCGCCCTTGGAGCGGA 5880  
Db 5953 TACCGCAGGGCGGACCGAGCTGCGCAGGGAAGAATCTCTGCGCCGCCCTTGGAGCGGA 6012  
QY 5881 GAGAAGTCCCGCGCGGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5940  
Db 6013 GAGAAGTCCCGCGCGGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGGCTGTTT 6072  
QY 5941 GAACACAGCAGCGGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGGTG 6000  
Db 6073 GAACACAGCAGCGGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGGTG 6132  
QY 6001 AACBAGGGAAGGGCA 6017  
Db 6133 AACBAGGGAAGGGCA 6149

RESULT 11  
US-10-028-946-3  
; Sequence 3, Application US/10028946  
; Publication No. US20020123622A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Miranda, Maricar  
; APPLICANT: Fiddle, Carl Johan  
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding t  
; FILE REFERENCE: LEX-0289-USA  
; CURRENT APPLICATION NUMBER: US/10/028,946  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: US 60/258,335  
; PRIOR FILING DATE: 2000-12-27



Db 1981 AATATCCGCCAGGCAAAAGGAGCGAGCGGAGAGCTGGAGAAGCTGCAGAACCGAGAG 2040  
QY 2011 GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGNACGCCGCCATCTCTG 2070  
Db 2041 GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGNACGCCGCCATCTCTG 2100  
QY 2071 GAGAACAAAGGTAAAGAGCTAGAGACCATGGAGCCGTAGAGAAACAGACTGAAGATGAC 2130  
Db 2101 GAGAACAAAGGTAAAGAGCTAGAGACCATGGAGCCGTAGAGAAACAGACTGAAGATGAC 2160  
QY 2131 ATCCAGACAAATCCCAACAGATCCAGACAGATGGCTGATATAATCTGAGACTCGAAGAG 2190  
Db 2161 ATCCAGACAAATCCCAACAGATCCAGACAGATGGCTGATATAATCTGAGACTCGAAGAG 2220  
QY 2191 AAACATCGGGAGGCCCAAGTCTCAGGCCAGCACCTAGAAAGTGCACCTGAAAACAGAAAGAG 2250  
Db 2221 AAACATCGGGAGGCCCAAGTCTCAGGCCAGCACCTAGAAAGTGCACCTGAAAACAGAAAGAG 2280  
QY 2251 CAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGCTGAC 2310  
Db 2281 CAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGCTGAC 2340  
QY 2311 AAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGCCCATGAGAAAGGCCAAA 2370  
Db 2341 AAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGCCCATGAGAAAGGCCAAA 2400  
QY 2371 ATTCTCAGCGAAACAGAGCGATGATCAATGTATGATTCCAAGATCAGATCCCTGGAA 2430  
Db 2401 ATTCTCAGCGAAACAGAGCGATGATCAATGTATGATTCCAAGATCAGATCCCTGGAA 2460  
QY 2431 CAGAGATTGTGGAATCTGAGCCCAATTAAGTGTGGCAATCAGATAAAGAAAGACCTGCTGAC 2490  
Db 2461 CAGAGATTGTGGAATCTGAGCCCAATTAAGTGTGGCAATCAGATAAAGAAAGACCTGCTGAC 2520  
QY 2491 CAAAGGAAATGAAGGCCCAAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550  
Db 2521 CAAAGGAAATGAAGGCCCAAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580  
QY 2551 CTGGAGACACAGCTGGGAAGTTGAGGCCCAAGAACCGGAACTGGAGAGCAGCTGAG 2610  
Db 2581 CTGGAGACACAGCTGGGAAGTTGAGGCCCAAGAACCGGAACTGGAGAGCAGCTGAG 2640  
QY 2611 AAGATCAGCCACCAAGACCACAGTACAGAAATCGGCTGCTGAACTGGAGACAAAGATTG 2670  
Db 2641 AAGATCAGCCACCAAGACCACAGTACAGAAATCGGCTGCTGAACTGGAGACAAAGATTG 2700  
QY 2671 CGGGAGGTCAGTCTAGACACGAGGACAGAACTGGAGCTCAAGCGCCAGCTCACAG 2730  
Db 2701 CGGGAGGTCAGTCTAGACACGAGGACAGAACTGGAGCTCAAGCGCCAGCTCACAG 2760  
QY 2731 CTACAGCTCTCCTCGAGAGCGGAGTCAAGTTCAGAGCCCTGAGGCTGCAGGGCG 2790  
Db 2761 CTACAGCTCTCCTCGAGAGCGGAGTCAAGTTCAGAGCCCTGAGGCTGCAGGGCG 2820  
QY 2791 GCCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAAGACCAACAGCAAGCT 2850  
Db 2821 GCCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAAGACCAACAGCAAGCT 2880  
QY 2851 GAAGAGGATCCAGGCACTCAGGCAATAGAGATGAAATCCAGCGCAATTTGATGCT 2910  
Db 2881 GAAGAGGATCCAGGCACTCAGGCAATAGAGATGAAATCCAGCGCAATTTGATGCT 2940  
QY 2911 CTTTCGTAAACAGTGTACTGTAATCAGAGCTGGAGGAGCAGTAAACAGCTGACCGAG 2970  
Db 2941 CTTTCGTAAACAGTGTACTGTAATCAGAGCTGGAGGAGCAGTAAACAGCTGACCGAG 3000  
QY 2971 GACACGCTGAATCAACAAACAAATTTCTACTTGTCCAAACAACTCGATGAGCTTCT 3030  
Db 3001 GACACGCTGAATCAACAAACAAATTTCTACTTGTCCAAACAACTCGATGAGCTTCT 3060  
QY 3031 GCGGCCACGACGAGATTGTACAACTGCGAGTGAAGTGGACCATCTCGCGGGAGATC 3090  
Db 3061 GCGGCCACGACGAGATTGTACAACTGCGAGTGAAGTGGACCATCTCTCGCGGGAGATC 3120

QY 3091 ACGGAACGAGAGATGCAGCTTACCAGCCAGAACAGATGGAGGCTCTGAAGACCACG 3150  
Db 3121 ACGGAACGAGAGATGCAGCTTACCAGCCAGAACAGATGGAGGCTCTGAAGACCACG 3180  
QY 3151 TGCACCATGTCTGGAGGAACAGGTCTATGGATTTGGAGGCCCTTAAACGATGAGCTCTGAA 3210  
Db 3181 TGCACCATGTCTGGAGGAACAGGTCTATGGATTTGGAGGCCCTTAAACGATGAGCTCTGAA 3240  
QY 3211 AAAAGCGGAGTGGAGGCTTGGAGAGCGTCTGGGTGATGAGAAATCCAGTTTGAG 3270  
Db 3241 AAAAGCGGAGTGGAGGCTTGGAGAGCGTCTGGGTGATGAGAAATCCAGTTTGAG 3300  
QY 3271 TGTCCGGTTCCAGAGCTTCGACAGAAATGCTGGACACCCAGAGAAACAGAGCAGGCCAGAGCC 3330  
Db 3301 TGTCCGGTTCCAGAGCTTCGACAGAAATGCTGGACACCCAGAGAAACAGAGCAGGCCAGAGCC 3360  
QY 3331 GATCAGCGGATCACCAGTCTGCCAGTGTGGAGTGGAGTGGAGTGAAGGAGCACAGGCT 3390  
Db 3361 GATCAGCGGATCACCAGTCTGCCAGTGTGGAGTGGAGTGGAGTGAAGGAGCACAGGCT 3420  
QY 3391 GAGATTCTCGCTCTGCAGCAGGCTCTCAAGAGCAGAGCTGAAAGCCGAGAGCTCTCT 3450  
Db 3421 GAGATTCTCGCTCTGCAGCAGGCTCTCAAGAGCAGAGCTGAAAGCCGAGAGCTCTCT 3480  
QY 3451 GACAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAAATGCCCGAAGCTTA 3510  
Db 3481 GACAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAAATGCCCGAAGCTTA 3540  
QY 3511 CAGCAGAGCTGGAGATGAAACGAGCTCAACAGAGGCTTCTGGAGAGCAGCCAAA 3570  
Db 3541 CAGCAGAGCTGGAGATGAAACGAGCTCAACAGAGGCTTCTGGAGAGCAGCCAAA 3600  
QY 3571 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCCGCTCTGACTCAAGGACTCAA 3630  
Db 3601 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCCGCTCTGACTCAAGGACTCAA 3660  
QY 3631 GAACTCTAGATCGGCTGATCTACTGAAGACAGAAAGATGACTTGGAGTATCAGCTG 3690  
Db 3661 GAACTCTAGATCGGCTGATCTACTGAAGACAGAAAGATGACTTGGAGTATCAGCTG 3720  
QY 3691 GAAACATTCAGGTTCTTCTTCTCATGAAAAGTGAATGGAAGGCACTATTTCTCAA 3750  
Db 3721 GAAACATTCAGGTTCTTCTTCTCATGAAAAGTGAATGGAAGGCACTATTTCTCAA 3780  
QY 3751 CAAACCAAACTATGATTTTCTGCAAGCCAAATGGAACCAACTGCTGTAAGAAAAAG 3810  
Db 3781 CAAACCAAACTATGATTTTCTGCAAGCCAAATGGAACCAACTGCTGTAAGAAAAAG 3838  
QY 3811 GGTTTATTTAGTCGACGAAAGAGGCCCTGCTTTACCCACACAGTTTCTCTGCACTAC 3870  
Db 3839 -----AGGTTCTCTGCACTAC 3855  
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTTAGAGGAGCC 3930  
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTTAGAGGAGCC 3915  
QY 3931 CTTTCAGAAACCCGATCGAGCTCCGTCGCCCGGAGGAAAGCTGCCACCCCAAGCA 3990  
Db 3916 CTTTCAGAAACCCGATCGAGCTCCGTCGCCCGGAGGAAAGCTGCCACCCCAAGCA 3975  
QY 3991 ACGGACCAACCAACCCATCCAGCCAGCCACCGCAGGAGCAGATCGCCATGTCGCC 4050  
Db 3976 ACGGACCAACCAACCCATCCAGCCAGCCACCGCAGGAGCAGATCGCCATGTCGCC 4035  
QY 4051 ATCGTCCGCTGCGCAGAGCAACCCAGCCAGTGCATGAGCTGTGSCCCCGCCATCCAGC 4110  
Db 4036 ATCGTCCGCTGCGCAGAGCAACCCAGCCAGTGCATGAGCTGTGSCCCCGCCATCCAGC 4095  
QY 4111 CGCAGAAAGAGTCTCAACTCCAGAGGATTTAGTCGGGCTCTTAAGGAACCCATGCGAC 4170  
Db 4096 CGCAGAAAGAGTCTCAACTCCAGAGGATTTAGTCGGGCTCTTAAGGAACCCATGCGAC 4155





QY 4945 GCCAGTCCACCTGCTGCTGCCAGCCCGACATCTCACCAACATTTTGAAGCTGTCAAG 5004  
Db 2176 GCCAGTCCACCTGCTGCCAGCCCGACATCTCACCAACATTTTGAAGCTGTCAAG 2235  
QY 5005 GGCTGCCACTTTGTTGGGCGAGCAAGATTGAGAACGGGCTCTGCATCTGTGACGCGATG 5064  
Db 2236 GGCTGCCACTTTGTTGGGCGAGCAAGATTGAGAACGGGCTCTGCATCTGTGACGCGATG 2295  
QY 5065 CCAGCAAGTGTCTATTTCCGCTCAACAGAAACCTCAGCAAAATACATGCAATCCGAAA 5124  
Db 2296 CCAGCAAGTGTCTATTTCCGCTCAACAGAAACCTCAGCAAAATACATGCAATCCGAAA 2355  
QY 5125 GAGATAGACCTCAGAGCCCTGCAGCTGTATCCACTTCCCAATACAGTATCTCATTT 5184  
Db 2356 GAGATAGACCTCAGAGCCCTGCAGCTGTATCCACTTCCCAATACAGTATCTCATTT 2415  
QY 5185 GGAACCAATTAATTTTACGAATTCGACATGAACGATGACGCTCAGCGAATTCCTGGAT 5244  
Db 2416 GGAACCAATTAATTTTACGAATTCGACATGAACGATGACGCTCAGCGAATTCCTGGAT 2475  
QY 5245 AAGATGACCATTCCTTGGCACCCTGCTGTGTTTGGCGCCTCTTCCAAACAGCTTCCCTGTC 5304  
Db 2476 AAGATGACCATTCCTTGGCACCCTGCTGTGTTTGGCGCCTCTTCCAAACAGCTTCCCTGTC 2535  
QY 5305 TCAATCGTCAGGTGAACAGCGGCGAGGAGGAGTACTTCTGCTGTGTTTCCAGAA 5364  
Db 2536 TCAATCGTCAGGTGAACAGCGGCGAGGAGGAGTACTTCTGCTGTGTTTCCAGAA 2595  
QY 5365 TTGAGAGTGTCTGATTTTACGGAACGCTAGCGGACAGAGGAGTACTTCAAGTGGAGT 5424  
Db 2596 TTGAGAGTGTCTGATTTTACGGAACGCTAGCGGACAGAGGAGTACTTCAAGTGGAGT 2655  
QY 5425 CGTTACCTTTGGCTTTGCTTACAGAGAACCTTATCTGTTTGTGACCCACTTCAACTCA 5484  
Db 2656 CGTTACCTTTGGCTTTGCTTACAGAGAACCTTATCTGTTTGTGACCCACTTCAACTCA 2715  
QY 5485 CTCGAAGTAAATGAGATCCAGGACGCTCTCAGCAGGAGACCCCTGCGGAGCGTACCTG 5544  
Db 2716 CTCGAAGTAAATGAGATCCAGGACGCTCTCAGCAGGAGACCCCTGCGGAGCGTACCTG 2775  
QY 5545 GACATCCGAACCCGCGCTACCTGGCCCTGCCATTTCTCAGGAGCGGATTTACTTGGCG 5604  
Db 2776 GACATCCGAACCCGCGCTACCTGGCCCTGCCATTTCTCAGGAGCGGATTTACTTGGCG 2835  
QY 5605 TCCTATACCGAGTAATTAAGGTCAATTCGTCAGGGAACCTCTGTGAAGGAGTCC 5664  
Db 2836 TCCTATACCGAGTAATTAAGGTCAATTCGTCAGGGAACCTCTGTGAAGGAGTCC 2895  
QY 5665 GGCACCTGAACACCCAGCGGCGCGCTCCACCTCCCGCAG 5702  
Db 2896 GGCACCTGAACACCCAGCGGCGCGCTCCACCTCCCGCAG 2933

## RESULT 13

US-10-262-511-7  
; Sequence 7, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:

; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C

; CURRENT APPLICATION NUMBER: US/10/262,511

; CURRENT FILING DATE: 2003-05-28

; PRIOR APPLICATION NUMBER: 60/326,483

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: 60/373,815

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/381,642

; PRIOR FILING DATE: 2002-05-17

; PRIOR APPLICATION NUMBER: 60/328,029

; PRIOR FILING DATE: 2002-10-09

; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-05-16

; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/373,260

; PRIOR FILING DATE: 2002-04-17

; PRIOR APPLICATION NUMBER: 60/373,826

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/327,435

; PRIOR FILING DATE: 2001-10-05

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 439

; SOFTWARE: CuroSeqlist version 0.1

; SEQ ID NO 7

; LENGTH: 2542

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (2)..(2542)

; US-10-262-511-7

Query Match 36.7%; Score 2414; DB 13; Length 2542;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 2475; Conservative 0; Mismatches 5; Indels 48; Gaps 1;

QY 1864 CAAGGGAAGCTTGAAGTGGAGAAATATCGAAACTGGAGAGATCAATCTGAGCAGCAG 1923

Db 11 CAGGGAAGCTTGAAGTGGAGAAATATCGAAACTGGAGAGATCAATCTGAGCAGCAG 70

QY 1924 CTCAAAATTCAGGAGCTCCCAAGAGAAACTCGGAGAGGCT----- 1962

Db 71 CTCAAAATTCAGGAGCTCCCAAGAGAAACTCGGAGAGGCTTAAAGCCAGCAGGAGCC 130

QY 1963 -----GCAAGAGCGAGCCGAGAGGAGCTTGGAGAG 1995

Db 131 ACCGAGCTGCTCAGAAATATCCGCCAGGCAAGAGCGAGCCGAGGAGGAGCTGGAGAG 190

QY 1996 CTCGAGAACCGAGAGGATTTCTTCTGAGGCAATCAGAAAGAGCTGGTGGAGCTGAGGAA 2055

Db 191 CTGCAAGAACCGAGAGGATTTCTTCTGAGGCAATCAGAAAGAGCTGGTGGAGCTGAGGAA 250

QY 2056 CGCCGCCATTTCTCTGGAGAACAAAGGTAAGAGACTAGAGACCATGAGCGGTAGAGAAAC 2115

Db 251 CGCCGCCATTTCTCTGGAGAACAAAGGTAAGAGACTAGAGACCATGAGCGGTAGAGAAAC 310



||||| 2471 GAATGTCAGTGTGTCACCCCAAGTGTCTCCACGTCCTCCAGCGCCACTCGCGCTTG 2530  
|||||  
||||| 4336 CCTGCTGA 4343  
|||||  
||||| 2531 CCTGTCGA 2538  
|||||

```

;
; NAME/KEY: CDS
; LOCATION: (2) .. (2497)
US-10-262-511-5

Query Match      35.2%; Score 2317.2; DB 13; Length 2497;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 2432; Conservative 0; Mismatches 3; Indels 93; Gaps 2;

QY 1864 CAAGGGAAGCCTGAAGTGGGAGAAATATCGGAAACTGGAGAGATCAATGCTGACGACAG 1923
Db 11 CAGGGGAAGCCTGAAGTGGGAGAAATATCGGAACTGGAGAGATCAATGCTGACGACAG 70
QY 1924 CTCAAATTCAGGAGCTCCAGAGAAACTGGAGAGGCT----- 1962
Db 71 CTCAAATTCAGGAGCTCCAGAGAAACTGGAGAGGCTGTAAAGCCAGACGAGAGGCC 130
QY 1963 -----GCAAAGGAGCGAGCCGAGAGGAGGAGGAGCTGGAGAG 1995
Db 131 ACCGAGCTGCTGCAGAAATATCCGCCAGGCAAGAGGAGCGCCGAGAGGAGGAGCTGGAGAG 190
QY 1996 CTGCAGAAACCGAGAGGATTTCTTGAAGGATCAGAAAGAGCTGGTGGAGCTGAGGAA 2055
Db 191 CTGCAGAAACCGAGAGGATTTCTTGAAGGATCAGAAAGAGCTGGTGGAGCTGAGGAA 250
QY 2056 CGCGCCATTTCTTGGAGAAACAAGTAAAGAGACTAGAGACCATGGAGCGCTAGAGAAAC 2115
Db 251 CGCGCCATTTCTTGGAGAAACAAGTAAAGAGACTAGAGACCATGGAGCGCTAGAGAAAC 310
QY 2116 AGACTGAAGGATGACATCCAGACAAAAATCCCAACAGATCCAGCAGATGGCTGATAAATT 2175
Db 311 AGACTGAAGGATGACATCCAGACAAAAATCCCAACAGATCCAGCAGATGGCTGATAAATT 370
QY 2176 CTGGAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTCAGCCCGAGCACTTAGAGTGCAC 2235
Db 371 CTGGAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTCAGCCCGAGCACTTAGAGTGCAC 430
QY 2236 CTGAAACAGAAAGAGCAGCAGCTATGAGGAAAGATTTAAAGTGTGGCAATCAGATAAG 2295
Db 431 CTGAAACAGAAAGAGCAGCAGCTATGAGGAAAGATTTAAAGTGTGGCAATCAGATAAG 490
QY 2296 AAAGACCTGGCTGACAGAGGAGACATCGAGAAACATGATGTCAGAGACACAGAGGAGGCGC 2355
Db 491 AAAGACCTGGCTGACAGAGGAGACATCGAGAAACATGATGTCAGAGACACAGAGGAGGCGC 550
QY 2356 CATGAGAAAGGCAAAATTTCTCAGCGAACAGAAAGCCGATGATCAATGCTATGGATTCCAAG 2415
Db 551 CATGAGAAAGGCAAAATTTCTCAGCGAACAGAAAGCCGATGATCAATGCTATGGATTCCAAG 610
QY 2416 ATCAGATCCCTGGAACAGAGGATTTGGAACTGTCTGAAAGCCAAATAAATTGAGCAAAAT 2475
Db 611 ATCAGATCCCTGGAACAGAGGATTTGGAACTGTCTGAAAGCCAAATAAATTGAGCAAAAT 670
QY 2476 AGCAGTCTTTTACCCAAAGGACATGAGGCCCAAGAGAGATGATTTCTGAACCTCAGG 2535
Db 671 AGCAGTCTTTTACCCAAAGGACATGAGGCCCAAGAGAGATGATTTCTGAACCTCAGG 730
QY 2536 CAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCGAAAACTG 2595
Db 731 CAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCGAAAACTG 790
QY 2596 GAGGAGCAGCTGGAGAGATCCACCAGAGACCAAGTACAGAGATCGGCTGCTGGAA 2655
Db 791 GAGGAGCAGCTGGAGAGATCCACCAGAGACCAAGTACAGAGATCGGCTGCTGGAA 850
QY 2656 CTGAGAGAAAGATTGGGGAGGTCAGTCTAGAGCAGCAGGAGCAGAAACTGGAGCTCAAG 2715
Db 851 CTGAGAGAAAGATTGGGGAGGTCAGTCTAGAGCAGCAGGAGCAGAAACTGGAGCTCAAG 910
QY 2716 CGCCAGCTCAGAGCTACAGCTCTCCCTGAGGAGCGGAGTTCAGTTGACAGCCCTG 2775
Db 911 CGCCAGCTCAGAGCTACAGCTCTCCCTGAGGAGCGGAGTTCAGTTGACAGCCCTG 970

RESULT 14
US-10-262-511-5
; Sequence 5, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kexuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Svytek, Kimberly A.
; APPLICANT: Edinger, Shomut R.
; APPLICANT: Eilerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; PRIOR FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 5
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
```

QY 2776 CAGGTGTCACGGCGGCGCTTGAGAGCCAGCTTCCGACGGGAGACAGACAGCTGCGAAG 2835  
Db 971 CAGGTGTCACGGCGGCGCTTGAGAGCCAGCTTCCGACGGGAGACAGACAGCTGCGAAG 1030  
QY 2836 ACCACAGCAGAGCTGAAGAGAGATCCAGGACATCCAGGACATAGAGATGAATCCAG 2895  
Db 1031 ACCACAGCAGAGCTGAAGAGAGATCCAGGACATCCAGGACATAGAGATGAATCCAG 1090  
QY 2896 GCGAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATCAGACCTCGAGGAGCAGCTA 2955  
Db 1091 GCGAATTTGATGCTCTTCTGTAACAGCTGTACTGTAATCAGACCTCGAGGAGCAGCTA 1150  
QY 2956 AACCCAGCTGACCGAGGACAAACGCTGAATCAACAAACCAAACTTCTACTTTGCCAAACA 3015  
Db 1151 AACCCAGCTGACCGAGGACAAACGCTGAATCAACAAACCAAACTTCTACTTTGCCAAACA 1210  
QY 3016 CTCGATGAGGCTTCTGGCGCCAAACGAGATGTACAACTCGGAAGTGAAGTGAACCAT 3075  
Db 1211 CTCGATGAGGCTTCTGGCGCCAAACGAGATGTACAACTCGGAAGTGAAGTGAACCAT 1270  
QY 3076 CTCGCGCGGGAGATCACGGAAACGAGATGCAAGCTTACCAGCCAGAAACCAACGATGGAG 3135  
Db 1271 CTCGCGCGGGAGATCACGGAAACGAGATGCAAGCTTACCAGCCAGAAACCAACGATGGAG 1330  
QY 3136 GCTCTGAAGACCAAGTGCACCATGCTGGAGGAAACAGGTCAAGATTTGGAGCCCTTAAC 3195  
Db 1331 GCTCTGAAGACCAAGTGCACCATGCTGGAGGAAACAGGTCAAGATTTGGAGCCCTTAAC 1390  
QY 3196 GATGAGCTCTAGAAAAGAGCGGAGTGGGAGGCTGGAGAGCGTCTGGGTGATGAG 3255  
Db 1391 GATGAGCTCTAGAAAAGAGCGGAGTGGGAGGCTGGAGAGCGTCTGGGTGATGAG 1450  
QY 3256 AAATCCAGCTTTGAGTGTCCGGTTCCAGAGCTGCAAGAGATCTGGACACCGAGAAAACAG 3315  
Db 1451 AAATCCAGCTTTGAGTGTCCGGTTCCAGAGCTGCAAGAGATCTGGACACCGAGAAAACAG 1510  
QY 3316 ACCAGGCGAGCCGATCAGCGGATCACCGAGTCTCGCAGGTGCGAGCTGCAGTG 3375  
Db 1511 ACCAGGCGAGCCGATCAGCGGATCACCGAGTCTCGCAGGTGCGAGCTGCAGTG 1570  
QY 3376 AAGGAGCAAGAGCTGAGATTTCTGCTGTCAGCAGGCTCTCAAAAGAGCAGAAAGCTGAAG 3435  
Db 1571 AAGGAGCAAGAGCTGAGATTTCTGCTGTCAGCAGGCTCTCAAAAGAGCAGAAAGCTGAAG 1630  
QY 3436 GCGAGAGCTCTCTGCAAGCTCAATGACCTGGAGAGAGCATGCTATGTTGAAATG 3495  
Db 1631 GCGAGAGCTCTCTGCAAGCTCAATGACCTGGAGAGAGCATGCTATGTTGAAATG 1690  
QY 3496 AATGCGCGAGCTTACAGCAGAGCTGGAGACTGAAACGAGAGCTCAAAAGAGGCTTCTG 3555  
Db 1691 AATGCGCGAGCTTACAGCAGAGCTGGAGACTGAAACGAGAGCTCAAAAGAGGCTTCTG 1750  
QY 3556 GAAGAGCAAGCCAAATTAACAGCAGCAGATGGACCTGCAGAAAATACATTTTCGCTG 3615  
Db 1751 GAAGAGCAAGCCAAATTAACAGCAGCAGATGGACCTGCAGAAAATACATTTTCGCTG 1810  
QY 3616 ACTCAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAGACAGAAAGAGTGCAC 3675  
Db 1811 ACTCAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAGACAGAAAGAGTGCAC 1870  
QY 3676 TTGGAGTATCAGCTGGAACCAATTCAGGTTCTCTATTCTCATGAAAAGCTGAAAATGGAA 3735  
Db 1871 TTGGAGTATCAGCTGGAACCAATTCAGGTTCTCTATTCTCATGAAAAGCTGAAAATGGAA 1930  
QY 3736 GGCATATTTCTCAACAAACCAATTCATGATTTCTGCAAGCCAAATGACCAACT 3795  
Db 1931 GGCATATTTCTCAACAAACCAATTCATGATTTCTGCAAGCCAAATGACCAACT 1990  
QY 3796 GCTAAAAGAAAAGGGTTTTATTAGTCGACGGAAGAGGACCTGCTTTACCCACACAG 3855  
Db 1991 GCTAAAAGAAAAGGGTTTTATTAGTCGACGGAAGAGGACCTGCTTTACCCACACAG 3915

Db 2006 GTTCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCCTCTGTGA 2065  
QY 3916 GAGCTAGAGAAAGCCCTTCAAGAGACCCGATCGAGTCCGGTCCGCCCGGAGGAAGCT 3975  
Db 2066 GAGCTAGAGAAAGCCCTTCAAGAGACCCGATCGAGTCCGGTCCGCCCGGAGGAAGCT 2125  
QY 3976 GCCCACCGCAAAAGCAACGGACCAACCCATCCAGCCAGCCACCGCGAGGACAG 4035  
Db 2126 GCCCACCGCAAAAGCAACGGACCAACCCATCCAGCCAGCCACCGCGAGGACAG 2185  
QY 4036 ATGCCATGTCGCCATCGTGGGTGCGAGACACCCAGCCAGTGCATGAGCCTGCTG 4095  
Db 2186 ATGCCATGTCGCCATCGTGGGTGCGAGACACCCAGCCAGTGCATGAGCCTGCTG 2245  
QY 4096 GCCCGCGCATCCAGCGCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCT 4155  
Db 2246 GCCCGCGCATCCAGCGCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCT 2305  
QY 4156 AAGAACGATGACACCAATTTCTCACCAGTCAACGTAGGACTGAACATGCGAGCC 4215  
Db 2306 AAGAACGATGACACCAATTTCTCACCAGTCAACGTAGGACTGAACATGCGAGCC 2365  
QY 4216 ACAAGTGTCTGTGTCTGTGATACCGTGCACCTTTGGACGCCAGGATCCAAATGCTC 4275  
Db 2366 ACAAGTGTCTGT 2425  
QY 4276 GAATGTGAGT 4335  
Db 2426 GAATGTGAGT 2485  
QY 4336 CCTGCTGA 4343  
Db 2486 CCTGCTGA 2493

RESULT 15  
US-10-262-511-3  
; Sequence 3, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Casterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511

```

; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Curasequid version 0.1
; SEQ ID NO 3
; LENGTH: 1870
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1870)
US-10-262-511-3

Query Match
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1849; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 6 CGCAGTGGGGAGATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGATGCTGGTGC 65
DB 1 CACCGGTACCACCATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGATGCTGGTGC 60
QY 66 TGTGTAACCATGTCAGCGCGGCTCCAGGTGATCTGTTCTTCCAGGGGAACCCACC 125
DB 61 TGTGTAAACCATGTCAGCGCGGCTCCAGGTGATCTGTTCTTCCAGGGGAACCCACC 120
QY 126 CTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAAGGGATATTAGATGCCCTCTT 185
DB 121 CTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAAGGGATATTAGATGCCCTCTT 180
QY 186 TGTCTCTTTGAGAAATGAGTCAGTCAGCTGCTCTGATGAAGATTAAGCAGTGAACAATT 245
DB 181 TGTCTCTTTGAGAAATGAGTCAGTCAGCTGCTCTGATGAAGATTAAGCAGTGAACAATT 240
QY 246 TGTCCGGAAGTATTCGACACCATAGCTGAGTTACAGAGCTCCAGCCCTTCGCAAGGA 305
DB 241 TGTCCGGAAGTATTCGACACCATAGCTGAGTTACAGAGCTCCAGCCCTTCGCAAGGA 300
QY 306 CTTCCGAGTCAAGATCTTTGAGTGTGGTGCATTTGCTGAAGTGAAGTGGTGAAGAGA 365
DB 301 CTTCCGAGTCAAGATCTTTGAGTGTGGTGCATTTGCTGAAGTGAAGTGGTGAAGAGA 360
QY 366 GAAAGCAACCGGGACATCTATGCTATGAAGTGAAGAGAGGCTTTTATGGCCCA 425
DB 361 GAAAGCAACCGGGACATCTATGCTATGAAGTGAAGAGAGGCTTTTATGGCCCA 420
QY 426 GGAGCAGGTTTCATTTTGGAGAGAGCGGAACATATTATCTCGAAGCAAGCCCGTG 485
DB 421 GGAGCAGGTTTCATTTTGGAGAGAGCGGAACATATTATCTCGAAGCAAGCCCGTG 480
QY 486 GATCCCCCAATTACAGTATGCCCTTTTCAGACAAAAATCACCTTTATCTGTGGAGGATA 545
DB 481 GATCCCCCAATTACAGTATGCCCTTTTCAGACAAAAATCACCTTTATCTGTGTGAGGATA 540

```

```

QY 546 TCAGCTGGAGGGAGCTTCTGCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAA 605
DB 541 TCAGCTGGAGGGAGCTTCTGCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAA 600
QY 606 CTTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTCCACGCGTTCATCTGATGGG 665
DB 601 CTTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTCCACGCGTTCATCTGATGGG 660
QY 666 ATACGTGATCGAGACATCAAGCCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAA 725
DB 661 ATACGTGATCGAGACATCAAGCCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAA 720
QY 726 GCTGGTGAATTTGATCTCGCGGGAATTAATTAACAAAGATGTTGAATGCAAACT 785
DB 721 GCTGGTGAATTTGATCTCGCGGGAATTAATTAACAAAGATGTTGAATGCAAACT 780
QY 786 CCCGATTGGGACCCCGAGATTACATGGCTTCTGAAAGTGTGACTGTGATGAACGGGGATGG 845
DB 781 CCCGATTGGGACCCCGAGATTACATGGCTTCTGAAAGTGTGACTGTGATGAACGGGGATGG 840
QY 846 AAAAGGCACCTACGGCCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 905
DB 841 AAAAGGCACCTACGGCCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 906 GATTTATGGGAGATCCCGCTTCGAGAGGGAACCTCTGCCAGAACCTTCCATTAACATTAT 965
DB 901 GATTTATGGGAGATCCCGCTTCGAGAGGGAACCTCTGCCAGAACCTTCCATTAACATTAT 960
QY 966 GAAATTCAGCGGTTTGTAAATTTCCAGATGACCCCAAGTGAAGTGAAGTTCCTTGA 1025
DB 961 GAAATTCAGCGGTTTGTAAATTTCCAGATGACCCCAAGTGAAGTGAAGTTCCTTGA 1020
QY 1026 TCTGATCAAAAGCTTGTGTCGCGCCAGAGAGAGACTGAAGTTTGAAGTCTTTGCTG 1085
DB 1021 TCTGATCAAAAGCTTGTGTCGCGCCAGAGAGAGACTGAAGTTTGAAGTCTTTGCTG 1080
QY 1086 CCATCTCTTCTCTAAAATTTGACTGGAACCAACATTCGTAACCTCTCTCCCCCTTCGT 1145
DB 1081 CCATCTCTTCTCTAAAATTTGACTGGAACCAACATTCGTAACCTCTCTCCCCCTTCGT 1140
QY 1146 TCCACACCTCAAGTCTGACGATGACACTCCAAATTTGATGAACAGAGAGAGATTCGCTG 1205
DB 1141 TCCACACCTCAAGTCTGACGATGACACTCCAAATTTGATGAACAGAGAGAGATTCGCTG 1200
QY 1206 GGTTCATCTCTCCGTGCGAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGT 1265
DB 1201 GGTTCATCTCTCCGTGCGAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGT 1260
QY 1266 TGTGGGTTTTGTAAGAGAGCACTGGGGATTTCTTGGTAGATCTGAGTCTGTTGTC 1325
DB 1261 TGTGGGTTTTGTAAGAGAGCACTGGGGATTTCTTGGTAGATCTGAGTCTGTTGTC 1320
QY 1326 GGGTCTGACTCCCTTCCCAAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAA 1385
DB 1321 GGGTCTGACTCCCTTCCCAAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAA 1380
QY 1386 AGAGCTACAAGATCTCAGGACAAAGTGTCAAGAGATGGAGAGAAATGACCCGGTTACA 1445
DB 1381 AGAGCTACAAGATCTCAGGACAAAGTGTCAAGAGATGGAGAGAAATGACCCGGTTACA 1440
QY 1446 TCCGAGAGTCTCAGAGTGGAGCTGCTGCTTAGTCAAGAGGAGGTGGAGCTGAGGCTTC 1505
DB 1441 TCCGAGAGTCTCAGAGTGGAGCTGCTGCTTAGTCAAGAGGAGGTGGAGCTGAGGCTTC 1500
QY 1506 TGAGACTCAGAGATCCCTCTGAGCAGGACCTTGTCTACCTACATCACAAGATGCAAGTAG 1565
DB 1501 TGAGACTCAGAGATCCCTCTGAGCAGGACCTTGTCTACCTACATCACAAGATGCAAGTAG 1560
QY 1566 CTTAAAGCGAAGTTTGGAGCAACCGATGGGTTGCCAGGAGGTGCAAGAGCACT 1625
DB 1561 CTTAAAGCGAAGTTTGGAGCAACCGATGGGTTGCCAGGAGGTGCAAGAGCACT 1620
QY 1626 GCAGCTTCTCCATGATATCAGAGAGCAGAGCGGAGCTTCCAGAAATCAAGAGCAGGA 1685

```

Db	1621	GCAGCTTCTCCATGATATCAGAGAGCAGCGGAGCTCCAGAAATCAAGAGCAGGA	1680
Qy	1686	GTACCAGGCTCAAGTGGAGAAATGAGGTTGATGAATCAGTTGGAAGAGATTTGT	1745
Db	1681	GTACCAGGCTCAAGTGGAGAAATGAGGTTGATGAATCAGTTGGAAGAGATCTTGT	1740
Qy	1746	CTCAGCAAGAAGACGGAGTGATCTCTACGAATCTGAGCTCAGAGAGTCTCGGCTTGC	1805
Db	1741	CTCAGCAAGAAGACGGAGTGATCTCTACGAATCTGAGCTCAGAGAGTCTCGGCTTGC	1800
Qy	1806	TGAAGATTTCAAGCGGAAGGCAGAGATCTCAGCATAAACTGTTGAAGCTAAGGATCA	1865
Db	1801	TGAAGATTTCAAGCGGAAGGCAGAGATCTCAGCATAAACTGTTGAAGCTAAGGATCA	1860
Qy	1866	AG 1867	
Db	1861	GG 1862	

Search completed: July 4, 2004, 16:15:14  
Job time : 2637 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:08:00 ; Search time 100 Seconds  
(without alignments)  
6477.589 Million cell updates/sec

Title: US-10-017-216-2

Perfect score: 10490  
Sequence: 1 MKKFKYARNFLDAGAEPI.....QLNGEIRQVQVEKSVLRIDYC 2053

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25.\*

1: sp\_archea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9890	94.3	2027	4	Q86UQ9
2	9656	92.0	2055	11	Q88938
3	7717.5	73.6	1641	11	Q88528
4	7521	71.7	1618	11	Q9CX19
5	4819	45.9	1032	11	Q8C1U1
6	2165.5	20.6	494	11	Q88937
7	2150.5	20.5	448	11	Q88527
8	1839.5	17.5	1854	5	Q9VTY8
9	1450	13.8	1732	11	Q54874
10	1421.5	13.6	1719	4	Q86XX3
11	1385.5	13.2	1760	4	Q9ULU5
12	1382.5	13.2	1711	4	Q9Y5S2
13	1373	13.1	1713	11	Q7TT49
14	1372	13.1	1702	11	O54875
15	1363	13.0	1713	11	Q7TT50
16	1353	12.9	1638	4	Q86XX2

#### ALIGNMENTS

#### RESULT 1

ID	Q86UQ9	PRELIMINARY;	PRT;	2027 AA.
AC	Q86UQ9;			
DT	01-JUN-2003 (TrEMBLrel. 24, Created)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Rho/rac-interacting citron kinase.			
GN	CIT.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Huang C.Q., Wu S.L., Shan Y.X., Liu S., Xiao P.J.;			
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY257469; AAP13528.1; -			
DR	GO; GO:0005622; C:intracellular; IEA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.			
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.			
DR	GO; GO:0005083; F:small GTPase regulatory/interacting protein. . ; IEA.			
DR	GO; GO:0007242; P:intracellular signaling cascade; IEA.			
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.			
DR	InterPro; IPR001180; Citron.			
DR	InterPro; IPR005479; CPase_L_D2.			
DR	InterPro; IPR001849; PH.			
DR	InterPro; IPR002219; DAG_PE-Bind.			
DR	InterPro; IPR000961; Pkinase_C.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR000861; REM_repeat.			
DR	InterPro; IPR002290; Ser_thr_pkinase.			
DR	InterPro; IPR008271; Ser_thr_pkin_AS.			
DR	InterPro; IPR001245; Tyr_pkinase.			
DR	Pfam; PF00780; CNH; 1.			
DR	Pfam; PF00130; DAG_PE-Bind; 1.			
DR	Pfam; PF00169; PH; 1.			
DR	Pfam; PF00069; pkinase; 1.			

Q81WQ7 homo sapien  
O01583 caenorhabdi  
P70335 mus musculu  
Q63644 rattus norv  
Q28021 bos taurus  
Q13464 homo sapien  
Q90Y37 brachydanio  
Q9UQN5 homo sapien  
O75116 homo sapien  
P70336 mus musculu  
O77819 oryctolagus  
Q73732 xenopus lae  
Q62868 rattus norv  
O44368 drosophila  
Q9U779 drosophila  
Q9VXE3 drosophila  
Q961D4 drosophila  
P92199 caenorhabdi  
Q98N6 gallus gall  
Q8AVM0 xenopus lae  
Q86XZ8 homo sapien  
Q99646 homo sapien  
O00565 homo sapien  
Q8SS10 encephalito  
Q98N5 gallus gall  
Q98S79 arabidopsis  
Q91W99 mesembryant

DR	Pfam: PF00433; pkinase C; 1.	
DR	ProDom: PD000001; ProC_kinase; 1.	
DR	SMART: SMC0109; Cl; 1.	
DR	SMART: SMC0036; CNH; 1.	
DR	SMART: SMC0233; PH; 1.	
DR	SMART: SMC0320; S TK; 1.	
DR	SMART: SMC0133; S TK X; 1.	
DR	SMART: SMC0219; TYR-KG; 1.	
DR	PROSITE: PS00867; CPSASE_2; 1.	
DR	PROSITE: PS00479; DAG_PE_BIND_DOM_1; 1.	
DR	PROSITE: PS00081; DAG_PE_BIND_DOM_2; 1.	
DR	PROSITE: PS50003; PH DOMAIN; 1.	
DR	PROSITE: PS00107; PROTEIN KINASE ATP; 1.	
DR	PROSITE: PS50011; PROTEIN KINASE DOM; 1.	
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; 1.	
KW	Kinase.	
SQ	SEQUENCE 2027 AA; 231429 MW; 6B1D8C3F661F357B CRC64;	
Query Match 94.3%; Score 9890; DB 4; Length 2027;		
Best Local Similarity 94.5%; Pred. No. 0;		
Matches 1954; Conservative 2; Mismatches 4; Indels 106; Gaps 3;		
QY	1 MLKFKYGARNPLDAGAAEPIASASRLNLPFOGKPPFMTQQQMSPLSREGILDALFVLFE 60	
DB	1 MLKFKYGARNPLDAGAAEPIASASRLNLPFOGKPPFMTQQQMSPLSREGILDALFVLFE 60	
QY	61 ECQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120	
DB	61 ECQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHFAEVQVVRKATG 120	
QY	121 DIYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDQXNHYLMEEYQPGG 180	
DB	121 DIYAMKMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYAFQDQXNHYLMEEYQPGG 180	
QY	181 DLASLNRYEDQDENLIFYLAELTAVHSVHLMGYVHRDIKPENILVDRTHGHIKLVD 240	
DB	181 DLASLNRYEDQDENLIFYLAELTAVHSVHLMGYVHRDIKPENILVDRTHGHIKLVD 240	
QY	241 GSAAKNKNMNAKLPIGTPDYMAPEVLTVMNGDGKGTGYLDCDWWSVGVIAYEM1YGR 300	
DB	241 GSAAKNKNMNAKLPIGTPDYMAPEVLTVMNGDGKGTGYLDCDWWSVGVIAYEM1YGR 300	
QY	301 SPFAEGTSARTFNIMNFORFLKFPDDPKVSSDFLDLIQSLCGQKERLKFGELCHHPFF 360	
DB	301 SPFAEGTSARTFNIMNFORFLKFPDDPKVSSDFLDLIQSLCGQKERLKFGELCHHPFF 360	
QY	361 SKIDMNNIRNSPPFPVFTLKSDDDTSNFDPEPEKNSVWSSPCQLSPSGFSGEELFPVGFS 420	
DB	361 SKIDMNNIRNSPPFPVFTLKSDDDTSNFDPEPEKNSVWSSPCQLSPSGFSGEELFPVGFS 420	
QY	421 YSKALGILGRSESVGLSDSPAKTSMEKKLIIKSKELQDSQDKCHKMEQEMTRLHRRVS 480	
DB	421 YSKALGILGRSESVGLSDSPAKTSMEKKLIIKSKELQDSQDKCHKMEQEMTRLHRRVS 480	
QY	481 EVEAVLSQKEVELKASFTQSLLEQDLATYITECSSLKRSRQARMEVSOEDDKALQLLH 540	
DB	481 EVEAVLSQKEVELKASFTQSLLEQDLATYITECSSLKRSRQARMEVSOEDDKALQLLH 540	
QY	541 DIPEQSRKLQEIKEQYQYQAEEMRLMMNLQLEDIVSARRRDLYESELRESRLAAEFK 600	
DB	541 DIPEQSRKLQEIKEQYQYQAEEMRLMMNLQLEDIVSARRRDLYESELRESRLAAEFK 600	
QY	601 RKATECHQKLLKADQCKEVEGEYAKLEKINAEQQLKIQELQKLEKAKVASTEATELLQ 660	
DB	601 RKATECHQKLLKADQCKEVEGEYAKLEKINAEQQLKIQELQKLEKAKVASTEATELLQ 660	
QY	649 --- AKERAARELEKIQNREDSSEGIRKKLVFAEPERRHSLENKVRLETMERENRLKDD 704	
DB	661 NIFQAKERAARELEKIQNREDSSEGIRKKLVFAE	
QY	705 IQTKSQIQQADKILIEKREKREAVSAQHLVHLKQKEQHYEBKIKVLDNQIKKDLAD 764	
DB	695 -----ELEKREKREAVSAQHLVHLKQKEQHYEBKIKVLDNQIKKDLAD 738	

QY	765 KETLENMORHEBEAHEKGLISEQKAMINAMDSKTRSLRQIVELSEANKLANSSLFT 824	
DB	739 KETLENMORHEBEAHEKGLISEQKAMINAMDSKTRSLRQIVELSEANKLANSSLFT 798	
QY	825 QRNMKAQEMISELRQOKFYLETQAGKLEAQNKRKLEBOLEKISHQDSDNRILLEETRL 884	
DB	799 QRNMKAQEMISELRQOKFYLETQAGKLEAQNKRKLEBOLEKISHQDSDNRILLEETRL 858	
QY	885 REVSLHEHEQKLELKRQLTQLSLQRESQLTALQAARALESQLRQAKTELEETATA 944	
DB	859 REVSLHEHEQKLELKRQLTQLSLQRESQLTALQAARALESQLRQAKTELEETATA 918	
QY	945 EEEIQALTARDEIQRKFDALRNSCTVITDLEQLNQLTEDNAELNNQNFYLSKQLEAS 1004	
DB	919 EEEIQALTARDEIQRKFDALRNSCTVITDLEQLNQLTEDNAELNNQNFYLSKQLEAS 978	
QY	1005 GANDEIVQLRSEVDHLRREITEREMQLTSQKOTMEALKTTCTMLEEQVMDLEALNDELLE 1064	
DB	979 GANDEIVQLRSEVDHLRREITEREMQLTSQKOTMEALKTTCTMLEEQVMDLEALNDELLE 1038	
QY	1065 KERQWEAWRSVLGDQKQFECRVRELQRLMDTEKQSRARADQDQITRESQVVELAVKEHA 1124	
DB	1039 KERQWEAWRSVLGDQKQFECRVRELQRLMDTEKQSRARADQDQITRESQVVELAVKEHA 1098	
QY	1125 EILALOQALKEQKLEKABESLSDKLNLEKHAMLEMMNARSLOQKLETERELKQRLLEQAK 1184	
DB	1099 EILALOQALKEQKLEKABESLSDKLNLEKHAMLEMMNARSLOQKLETERELKQRLLEQAK 1158	
QY	1185 LQQQMDLOKNHIFRLITQGLQALDRDLKLTESDLEYQLENIQVLSHEKVKMEGTISQ 1244	
DB	1159 LQQQMDLOKNHIFRLITQGLQALDRDLKLTESDLEYQLENIQVLSHEKVKMEGTISQ 1218	
QY	1245 QTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALTPQVPLQYNEKLKALEKARCASLEEA 1304	
DB	1219 QTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALTPQVPLQYNEKLKALEKARCASLEEA 1278	
QY	1305 LQKTRIELRSARBEAHRKATDHPHSTPTATQQAQAMSAIVRSPHQPMSAMSLAPPSS 1364	
DB	1279 LQKTRIELRSARBEAHRKATDHPHSTPTATQQAQAMSAIVRSPHQPMSAMSLAPPSS 1338	
QY	1365 RRKESSTPEFSRRLKERMHNTPHRFNVLNMRATKCAVCLDTVHFGROASKCLSCQVM 1424	
DB	1339 RRKESSTPEFSRRLKERMHNTPHRFNVLNMRATKCAVCLDTVHFGROASKCLSCQVM 1398	
QY	1425 CHPKCSTCLPATCGLPABYATHTEAFCRDKMNSPGLQTKPESSSLHLEGMKVPNNKR 1484	
DB	1399 CHPKCSTCLPATCGLPABYATHTEAFCRDKMNSPGLQTKPESSSLHLEGMKVPNNKR 1458	
QY	1485 GQQGWDKRYIVLGSKVLIVDNEAREAGQRPVEFELCLPDGDVSIHGAVGASELANTAK 1544	
DB	1459 GQQGWDKRYIVLGSKVLIVDNEAREAGQRPVEFELCLPDGDVSIHGAVGASELANTAK 1518	
QY	1545 A-----EKAEDAKL 1554	
DB	1519 ADVPYILKMSHPHTTCWPQRTLYLLAPSPDQKRWVTALESVVAGRVSRKAEADAKL 1578	
QY	1555 LGNSLLKLEGDDRLDNKCTLPSDQVVLVGTESGLYALNVLKNSLTHVPGIGAVFOIYII 1614	
DB	1579 LGNSLLKLEGDDRLDNKCTLPSDQVVLVGTESGLYALNVLKNSLTHVPGIGAVFOIYII 1638	
QY	1615 KDLEKLMIAAGEBRALCLVDVKKVKQSLAQSHLPAQDPDISPNIFEAVKGCFLFGAGKIEN 1674	
DB	1639 KDLEKLMIAAGEBRALCLVDVKKVKQSLAQSHLPAQDPDISPNIFEAVKGCFLFGAGKIEN 1698	
QY	1675 GLCICAMPKSKVILRYNENLSKYCTRKEIETSEPCSCIFHTNYSILITGNKFEYDMQ 1734	
DB	1699 GLCICAMPKSKVILRYNENLSKYCTRKEIETSEPCSCIFHTNYSILITGNKFEYDMQ 1758	
QY	1735 YTLEEFLLKNDHSLAPAVFAASSNSFPVSIQVNSAGQREYLLCFHFEGFVDSYGRS 1794	
DB	1759 YTLEEFLLKNDHSLAPAVFAASSNSFPVSIQVNSAGQREYLLCFHFEGFVDSYGRS 1818	

QY 1795 RTDCLKWSRLPLAFAYREPYLFVTHNSLEVIEIQARSSAGTAPARAYLDIPNPRYLGPAL 1854  
DB 1819 RTDCLKWSRLPLAFAYREPYLFVTHNSLEVIEIQARSSAGTAPARAYLDIPNPRYLGPAL 1878  
QY 1855 SSGAIYASSYQDKLVIKCKGNLVKESGTEHHRGPTSRSPKNGPPTYNHITKVA 1914  
DB 1879 SSGAIYASSYQDKLVIKCKGNLVKESGTEHHRGPTSRSPKNGPPTYNHITKVA 1938  
QY 1915 SSPAPPEGSPHREPSTPHRYREGTRTELKRDKSPGRPLERKSPGRMLSTRRERSPGRLF 1974  
DB 1939 SSPAPPEGSPHREPSTPHRYREGTRTELKRDKSPGRPLERKSPGRMLSTRRERSPGRLF 1998  
QY 1975 EDSSGRLLPAGAVRTPLSQVKNKGQSA 2002  
DB 1999 EDSSGRLLPAGAVRTPLSQVKNKGQSS 2026

RESULT 2  
ID O88938  
AC O88938 PRELIMINARY; PRT; 2055 AA.  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Rho/rac-interacting citron kinase.  
GN CIT OR CRK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=99009084; PubMed=9792683;  
RA Di Cunto F., Calautti E., Hsiao J., Ong L., Topley G., Turco E.,  
RA Dotto G.P.;  
RT "Citron rho-interacting kinase, a novel tissue-specific Ser/Thr kinase  
encompassing the rho-rac-binding protein citron.";  
J. Biol. Chem. 273:29706-29711(1998).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.  
EMBL; AF068624; AAC72823.1; -;  
DR HSP; P00517; I5TC.  
DR MGD; MGI:105313; Cit.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein.; IEA.  
DR GO; GO:0006740; F:transferrase activity; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR01180; Citron.  
DR InterPro; IPR005479; CCase L D2.  
DR InterPro; IPR002219; DAG\_PE-Bind.  
DR InterPro; IPR001849; PH.  
DR InterPro; IPR000961; Pkinase C.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR000861; REM repeat.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR InterPro; IPR000533; Tropomyosin.  
DR Pfam; PF00780; CNH; 1.  
DR Pfam; PF00130; DAG\_PE-bind; 1.  
DR Pfam; PF00169; PH; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00433; pkinase C; 1.  
DR PRINTS; PR00194; TROPOMYOSIN.  
DR PRODOM; PD000001; Prot\_kinase; 1.  
DR SMART; SM00109; Cl; 1.  
DR SMART; SM00036; CNH; 1.  
DR SMART; SM00233; PH; 1.  
DR SMART; SM00220; S\_TK; 1.  
DR SMART; SM00133; S\_TK\_X; 1.

DR PROSITE; PS00867; CPSASE 2; 1.  
DR PROSITE; PS00479; DAG\_PE\_BIND\_DOM 1; 1.  
DR PROSITE; PS00081; DAG\_PE\_BIND\_2; 1.  
DR PROSITE; PS00003; PH\_DOMAIN; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 2055 AA; 235481 MW; 2120CB5E454DA940 CRC64;  
Query Match 92.0%; Score 9656; DB 11; Length 2055;  
Best Local Similarity 91.94; Pred. No. 0;  
Matches 1903; Conservative 41; Mismatches 42; Indels 84; Gaps 5;  
QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFWTQQMSPLRGEGILDALFALFE 60  
DB 1 MLKFKYGVNPPEASASEPIASRASRLNLFQGGKPPFWTQQMSALSRGMLDALFALFE 60  
QY 61 ECSOPALMKIKHVNFRVKYSDTTIAELQELQPSAKDFEVSRLVCCGHFAEVQVVRKATG 120  
DB 61 ECSOPALMKHVSFFVKYSDTTIAELRELQPSARDFEVSRLVCCGHFAEVQVVRKATG 120  
QY 121 DIYAMKMKKALLAQOVQVFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYQPGG 180  
DB 121 DVIYAMKMKKALLAQOVQVFFEEERNILSRSTSPWIPQLQYAFQDKNHLVLMEEYQPGG 180  
QY 181 DLLSLNRYSDQDENLIQFYLAELILAVSHVLMGVYHRDIKPNILVDRTGHIKLVD 240  
DB 181 DFLSLNRYSDQDENSIQFYLAELILAVSHVLMGVYHRDIKPNILVDRTGHIKLVD 240  
QY 241 GSAAKMNSKMNNAKLPIGTDPYNAPEVLTVMAGDGKGTGCLDCDMSVGVIAEMILYGR 300  
DB 241 GSAAKMNSNK-VDAKLPITDPYNAPEVLTVMAGDRGTGCLDCDMSVGVIAEMIVYK 299  
QY 301 SPFAEGTSARTFNNIMNFORFLFPDDPKVSSDFDLIOQLLCOQKRLKFEGLCHCHPFF 360  
DB 300 TPFTGTSARTFNNIMNFORFLFPDDPKVSSSELDLQLSLLCVQKRLAFEGLCCHPFF 359  
QY 361 SKTDWNIRNSPPFPVTLKSDDDTSNFDPEPKNSWSSPCQLSPSGFSGEELPFVGF 420  
DB 360 ARTDWNIRNSPPFPVTLKSDDDTSNFDPEPKNSWAFILCVPAEPLAFEGELPFVGF 419  
QY 421 YSKALGILGSESVWSGLDSPAKTSMEKKLLIKSKELQSDQKCHQWEMRLHRRVS 480  
DB 420 YSKALGYLGSESVWSGLDSPAKVSSMEKKLLIKSKELQSDQKCHQWEMRLHRRVS 479  
QY 481 EYEAVALSQKVELKASQTORSLLLEQDLATYITCSSLKRSLEQARMEVSGEDDKALQLLH 540  
DB 480 EYEAVALSQKVELKASQTORSLLLEQDLATYITCSSLKRSLEQARMEVSGEDDKALQLLH 539  
QY 541 DIREQSKLOEIKEOEQYQAOVEEMRLMMNOLBEDLVSARRSDLYSELSERSESLAAEPEK 600  
DB 540 DIREQSKLOEIKEOEQYQAOVEEMRLMMNOLBEDLVSARRSDLYSELSERSESLAAEPEK 599  
QY 601 RKATECOHKLKAKDQKPEVGEYAKLEKINAEQQLKIQLQEKLEKA----- 648  
DB 600 RKANECQHKLMKAKDQKPEVGEYSKLEKINAEQQLKIQLQEKLEKAVKASATEATELLQ 659  
QY 649 ----AKERAERLEKLNREDSDSEGTRKLVAEERHSLKNVKELETWERENLKKD 704  
DB 660 NIRQAXERARELEKLNREDSDSEGTRKLVAEERHSLKNVKELETWERENLKKD 719  
QY 705 IQTKSQIQQAMADKILEBEKHREAOVSQAHLEVLKQKEQHYEEKIKVLDNQIKKDLAD 764  
DB 720 IQTKSEIQQAMADKILEBEKHREAOVSQAHLEVLKQKEQHYEEKIKVLDNQIKKDLAD 779  
QY 765 KETLENMQHREBEAHEKGIKILSEQAMINAMDSKIRSLERQIVELSEANKLAANSSFLT 824  
DB 780 KESLENMQHREBEAHEKGIKILSEQAMINAMDSKIRSLERQIVELSEANKLAANSSFLT 839  
QY 825 QRNKAQEEWISSELROOKFYLETQACKLAQAKNKELEQLEKISHQDHSKRNLELETRL 884  
DB 840 QRNKAQEEWISSELROOKFYLETQACKLAQAKNKELEQLEKISHQDHSKRNLELETRL 899

QY	885	REVSLEHERQKLEKQLTELQISLOERESQITAIQARAALESQIQAQTELEBTTAEA	944
DB	900	REVSLEHEEQKLEKQLTELQISLOERESQITAIQARAALESQIQAQTELEBTTAEA	959
QY	945	EIEIQALTAHREIQKFDALNSCTVITDLEEQNLQNTEDNAELNNQNFYLSKQLDEAS	1004
DB	960	EIEIQALTAHREIQKFDALNSCTVITDLEEQNLQNTEDNAELNNQNFYLSKQLDEAS	1019
QY	1005	GANDEIVQLRSEVDHLRRIITRETVOLTSQKOTMELKTKTTCMLBEQVMDLEALNDELLE	1064
DB	1020	GANDEIVQLRSEVDHLRRIITRETVOLTSQKOTMELKTKTTCMLBEQVMDLEALNDELLE	1079
QY	1065	KEQWEAWSVLGDEKSEFECVRELORMLDTEKOSRADORITTESROVVELAVKEHKA	1124
DB	1080	KEQWEAWSVLGDEKSEFECVRELORMLDTEKOSRADORITTESROVVELAVKEHKA	1139
QY	1125	EILALQALKEQKLAESLSDKLNLEKXHAMLENMARSLOQKLETERELKORLLEBOAK	1184
DB	1140	EILALQALKEQKLAESLSDKLNLEKXHAMLENMARSLOQKLETERELKORLLEBOAK	1199
QY	1185	LOQOMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYOLENTQVLYSHEKVMGTTISQ	1244
DB	1200	LOQOMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYOLENTQVLYSHEKVMGTTISQ	1259
QY	1245	QTKLIDFLQAKMDQPAKKGKLFRRKEDPALPTQVPLQYNELKLALEKEKARCAELSEA	1304
DB	1260	QTKLIDFLQAKMDQPAKKGKLFRRKEDPALPTQVPLQYNELKLALEKEKARCAELSEA	1304
QY	1305	LQKTRIELRSAREAAHRAKATDHPHSTPATARQIAMSIVRSPEHPSANSLAPSS	1364
DB	1305	LQKTRIELRSAREAAHRAKATDHPHSTPATARQIAMSIVRSPEHPSANSLAPSS	1364
QY	1365	RKESSTPEFRRRLKERMHNTPHFNVLNMRATKCAVCLDTHVFGQASKLEECQVM	1424
DB	1365	RKESSTPEFRRRLKERMHNTPHFNVLNMRATKCAVCLDTHVFGQASKLEECQVM	1424
QY	1425	CHPKCSTCLPATCGLPAEYATHEAFCDKXNSPGLQKPESSSLHLEGWKVPNNKR	1484
DB	1425	CHPKCSTCLPATCGLPAEYATHEAFCDKXNSPGLQKPESSSLHLEGWKVPNNKR	1484
QY	1485	GOQGWDRKYIVLGSKVLIYDNEAREAGORPVEFELCLPDGDVSIHGAVGASELANTAK	1544
DB	1485	GOQGWDRKYIVLGSKVLIYDNEAREAGORPVEFELCLPDGDVSIHGAVGASELANTAK	1544
QY	1545	A-----EKAEADAKL	1554
DB	1545	ADVPYILKMBHPHTTCWPORTLYLLAPSPDKQWVTALESVVAGGRVSREKAEADAKL	1604
QY	1555	LGNSLLKLEGGDRLDNMCTLPFSDQVVLVCTEGELVALNVLKNSLTHVPGIGAVFOIYII	1614
DB	1605	LGNSLLKLEGGDRLDNMCTLPFSDQVVLVCTEGELVALNVLKNSLTHVPGIGAVFOIYII	1664
QY	1615	KDLEKLMIAGERALCLVDVKKVKOSLAQSHLPAQPDISPNIPEAVKGCFLFGACKIEN	1674
DB	1665	KDLEKLMIAGERALCLVDVKKVKOSLAQSHLPAQPDISPNIPEAVKGCFLFGACKIEN	1724
QY	1675	GLCICAMPKSVILRNENLSKYIKKEITETSEPCSCIHFTNYSILIGTKNFYEIDMKQ	1734
DB	1725	SLCICAMPKSVILRNENLSKYIKKEITETSEPCSCIHFTNYSILIGTKNFYEIDMKQ	1784
QY	1735	YTLDEFLDKNDHSLAPAVFAASSNPFVSIQVNSAGQREYLLCFHEFGVFDVSYGRS	1794
DB	1785	YTLDEFLDKNDHSLAPAVFAASSNPFVSIQVNSAGQREYLLCFHEFGVFDVSYGRS	1844
QY	1795	RTDCLKWSRLPLAFAYREPYLVFTHFNLSVIEIHIQAPSSAGTTPARAYLDIPNRYLGPAL	1854
DB	1845	RTDCLKWSRLPLAFAYREPYLVFTHFNLSVIEIHIQAPSSAGTTPARAYLDIPNRYLGPAL	1904
QY	1855	SSGAIYVLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSSSNKXGPPYNEHITKRYA	1914
DB	1905	SSGAIYVLASSYQDKLRVICCKGNLVKESGTEHHRGPSTSSSNKXGPPYNEHITKRYA	1964

QY	1915	SSPAPPEGSHPREPSTPHRY--REGRTELRRDKSPGRPLERKSPGRMLSTRERSPGR	1972
DB	1965	SSPAPPEGSHPREPSTPHRYRDRCKRTELRRDKSPGRPLERKSPGRMLSTRERSPGR	2024
QY	1973	LFEDSSGRPLPAGAVRTPLSOVNKGQSA	2002
DB	2025	LFEDSSGRPLPAGAVRTPLSOVNKGQSS	2054
RESULT 3			
QY	088528	PRELIMINARY;	PRT; 1641 AA.
AC	088528;		
DT	01-NOV-1998	(TREMBlrel. 08, Created)	
DT	01-NOV-1998	(TREMBlrel. 08, Last sequence update)	
DT	01-OCT-2003	(TREMBlrel. 25, Last annotation update)	
DE	Citron-K kinase (Fragment).		
GN	CIT.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISUS=Brain.		
RX	MEDLINE=96128238; PubMed=8543060;		
RA	Madaule P., Furiyashiki T., Reid T., Ishizaki T., Watanabe G.,		
RA	Morii N., Narumiya S.;		
RT	"A novel partner for the GTP-bound forms of rho and rac.";		
RL	FEBS Lett. 377:243-248(1995).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISUS=Brain;		
RA	Madaule P., Eda M., Watanabe N., Fujisawa K., Matsuoka T., Bito H.,		
RA	Ishizaki T., Narumiya S.;		
RT	"Role of Citron kinase as a target of the small GTPase Rho in		
RT	cytokinesis.";		
RL	Nature 0:0-0(1998).		
CC	-1. SIMILARITY: CONTAINS 1 PH DOMAIN.		
DR	EMBL; AF070066; AAC27933.1;		
DR	MGI; MGI:105313; Cit.		
DR	GO; GO:0005622; C:intracellular; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0016301; F:kinase activity; IEA.		
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.		
DR	GO; GO:0005083; F:small GTPase regulatory/interacting protein. . ; IEA.		
DR	GO; GO:0007242; P:intracellular signaling cascade; IEA.		
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
DR	InterPro; IPR001180; Citron.		
DR	InterPro; IPR005479; Cbase L_D2.		
DR	InterPro; IPR002219; DAG PE-bind.		
DR	InterPro; IPR001849; PH.		
DR	InterPro; IPR000961; Pkinase C.		
DR	InterPro; IPR000861; REM_repeat.		
DR	Pfam; PF00780; CNH; 1.		
DR	Pfam; PF00130; DAG PE-bind; 1.		
DR	Pfam; PF00169; PH; 1.		
DR	SMART; SM00109; C1; 1.		
DR	SMART; SM00036; CNH; 1.		
DR	SMART; SM00233; PH; 1.		
DR	SMART; SM00133; S-TK X; 1.		
DR	PROSITE; PS00867; CFSASE_2; 1.		
DR	PROSITE; PS00479; DAG PE_BIND_DOM_1; 1.		
DR	PROSITE; PS00081; DAG PE_BIND_DOM_2; 1.		
DR	PROSITE; PS00003; PH_DOMAIN; 1.		
KW	Kinase.		
FT	NON_TER		
QY	1	187659 MW; 6723E20ECA3F22AB CRC64;	
SEQ	SEQUENCE	1641 AA;	
Query Match 73.6%; Score 7717.5; DB 11; Length 1641;			
Best Local Similarity 90.8%; Pred. No. 3.2e-316;			
Matches 1541; Conservative 15; Mismatches 16; Indels 125; Gaps 5;			

QY	374	PFVPTLKSDDTSNFTPEPKNSWVSSPCQLSPGSGBELPFVGFYSKALGILGRSES	433
Db	1	PFVPTLKSDDTSNFTPEPKNSWVSSVQLSPGSGBELPFVGFYSKALGILGRSES	60
QY	434	VVSGLDSPAKTSMEXKLLIKSKELQSDQKCHQWEOENTRLHRRVSEVAVLSQKEVEL	493
Db	61	VVSSLDSPAKVSMEXKLLIKSKELQSQKCHQWEOENTRLHRRVSEVAVLSQKEVEL	120
QY	494	KASQTSRLLEQDLATYITCSSLKRSLEQARMEVSQEDDKALQLLHDIREQSRKLQEI	553
Db	121	KASQTSRLLEQDLATYITCSSLKRSLEQARMEVSQEDDKALQLLHDIREQSRKLQEI	180
QY	554	EOYQAOQVEEMLMNQLBEDIVSARRSDLYSELRSLAABEFKRYKATCQHKLLKA	613
Db	181	EOYQAOQVEEMLMNQLBEDIVSARRSDLYSELRSLAABEFKRYKATCQHKLLKA	240
QY	614	KQQGEVGEYAKLEKINAEQQLKIQELOEKLEKA-----AKERAEREL	657
Db	241	KQQGEVGEYSKLEKINAEQQLKIQELOEKLEKAVKASTATELLQNIROKERAEREL	300
QY	658	EKLQNRDSESGIKKLVAABERRHSLNKKVRLTMRERENRLKDDIQTSQQIQQVAD	717
Db	301	EKLHNRDSESGIKKLVAE-----	321
QY	718	KILELEKHREQAQVSAQHLVHLKQEQHYEBKIKVLDNQIKKDLADKETTLENNMORHEE	777
Db	322	---ELEEKHREQAQVSAQHLVHLKQEQHYEBKIKVLDNQIKKDLADKETTLENNMORHEE	378
QY	778	EAHEKGKILSEQAMINAMDSKIRLEQRIVELSEANKLAANSLSFTORNKQAQEMISE	837
Db	379	EAHEKGKILSEQAMINAMDSKIRLEQRIVELSEANKLAANSLSFTORNKQAQEMISE	438
QY	838	LFQOQFYLETQAGKLEAQNKLQLEQLEKISHODHDKGRLLLETRLEVSLEHBEQKLE	897
Db	439	LFQOQFYLETQAGKLEAQNKLQLEQLEKISHODHDKGRLLLETRLEVSLEHBEQKLE	498
QY	898	LKRQLTQLSLQERESQLTALQAARAALESQRLQAKTELETTAAEBEETQALTAHRDE	957
Db	499	LKRQLTQLSLQERESQLTALQAARAALESQRLQAKTELETTAAEBEETQALTAHRDE	558
QY	958	IQRKFDALNSCTVITDLBEQNLQNTEDNAELNNQNFYLSKOLDASGANDEIVOLRSEV	1017
Db	559	IQRKFDALNSCTVITDLBEQNLQNTEDNAELNNQNFYLSKOLDASGANDEIVOLRSEV	618
QY	1018	DHLRREITERENQLTQSQOTMEALKTCTMLBEQVMDLEALNDELLEKEROWEAWRSVLG	1077
Db	619	DHLRREITERENQLTQSQOTMEALKTCTMLBEQVMDLEALNDELLEKEROWEAWRSVLG	678
QY	1078	DEKSOFEQVRELOQMLDTEKQSRARADQRTESQVVELAVKEHKABILAQALKEQK	1137
Db	679	DEKSOFEQVRELOQMLDTEKQSRARADQRTESQVVELAVKEHKABILAQALKEQK	738
QY	1138	LKAEISDKINDLEKXHAMLENNAISLOQKLETERELKORLLEEQAKLOQOQMDLQKNHIF	1197
Db	739	LKAEISDKINDLEKXHAMLENNAISLOQKLETERELKORLLEEQAKLOQOQMDLQKNHIF	798
QY	1198	RLTQGLQALDRADLLKTERSLEYQLENIQVLYSHEKVKMEGTISQQTKLIDFLQAKMD	1257
Db	799	RLTQGLQALDRADLLKTERSLEYQLENIQVLYSHEKVKMEGTISQQTKLIDFLQAKMD	858
QY	1258	QPAKKKKGLFSRKEDPALPTQVPIQYNELKLALEKEKARCAEALQKTRIELRSARE	1317
Db	859	QPAKKKK-----VPIQYNELKLALEKEKARCAEALQKTRIELRSARE	903
QY	1318	EAAHKKATDHPSTPATARQOIAMSAIVRSEHPQPSAMSLAPPSSRRKESSTPEFSR	1377
Db	904	EAAHKKATDHPSTPATARQOIAMSAIVRSEHPQPSAMSLAPPSSRRKESSTPEFSR	963
QY	1378	RLKERMHNIPIHFRNVGLNMRATKCAVCLDTHVFGQASKCLEQVCMCHPCSTCLPATC	1437
Db	964	RLKERMHNIPIHFRNVGLNMRATKCAVCLDTHVFGQASKCLEQVCMCHPCSTCLPATC	1023
QY	1438	GLPAEYATHFTEAFCDKMNKSPGLQTKPESSLLHSGWMMKVRNNKRGQQQWDRKYIVLE	1497

RESULT 4

Q9QX19	ID	Q9QX19	PRELIMINARY;	PRT;	1618	AA.
AC	Q9QX19;					
DT	01-MAY-2000	(TrEMBLrel. 13, Created)				
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)				
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)				
DE	Postsynaptic density protein.					
GN	CITRON.					
OS	Rattus norvegicus (Rat).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
OX	NCBI TaxID=10116;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RA	Zhang W., Apperson M.L., Vasquez L.E., Kennedy M.B.;					
RT	"Citron, a PSD-95-binding protein at glutamatergic synapses on					
RT	inhibitory neurons."					
RL	Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.					
CC	-1- SIMILARITY: CONTAINS 1 PH DOMAIN.					
DR	EMBL: AF039218; AAC25483.1; ..					
DR	GO: GO:0005822; C:intracellular; IEA.					
DR	GO: GO:0005524; F:ATP binding; IEA.					
DR	GO: GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.					
DR	GO: GO:0007242; P:intracellular signaling cascade; IEA.					
DR	InterPro: IPR001180; Citron.					
DR	InterPro: IPR005479; CPhase L D2.					
DR	InterPro: IPR002219; DAG_P5_Sind.					
DR	InterPro: IPR001849; PH.					

InterPro: IPR000861; REM repeat.  
 DR InterPro: IPR000533; Tropomyosin.  
 DR Pfam: PF00780; CNH; 1.  
 DR Pfam: PF00130; DAG\_PE\_bind; 1.  
 DR Pfam: PF00169; PH; 1.  
 DR PRINTS: PRO0194; TROPOMYOSIN.  
 DR SMART: SM00109; CI; 1.  
 DR SMART: SM00036; CNH; 1.  
 DR SMART: SM00233; PH; 1.  
 DR PROSITE: PS00867; CPASE 2; 1.  
 DR PROSITE: PS00479; DAG\_PE\_BIND\_DOM\_1; 1.  
 DR PROSITE: PS00881; DAG\_PE\_BIND\_DOM\_2; 1.  
 DR PROSITE: PS00003; PH\_DOMAIN; 1.  
 DR SEQUENCE 1618 AA; 185925 MW; EAED8073788AA429 CRC64;  
 Query Match 71.7%; Score 7521; DB 11; Length 1618;  
 Best Local Similarity 92.7%; Pred. No. 5.8e-308;  
 Matches 1500; Conservative 18; Mismatches 16; Indels 84; Gaps 5;  
 QY 468 MEQMTLHRRVSEPAVLISQKVELKASGTORSLLQDLATYITECSLSKRSLEQARME 527  
 DB 1 MEQEMARLHRRVSEPAVLISQKVELKASGTORSLLQDLATYITECSLSKRSLEQARME 60  
 QY 528 VSGEDDKAQLLDHIREQSKQEIKEQYQAVQVEEMRLMNMQLDEDLVARRRDLYES 587  
 DB 61 VSGEDDKAQLLDHIREQSKQEIKEQYQAVQVEEMRLMNMQLDEDLVARRRDLYES 120  
 QY 588 ELRESLAAEFPRKAKTECOHKLKAKDQKPEVGEYAKLEKINAEQQLKIQELQEKLEK 647  
 DB 121 ELRESLAAEFPRKAKTECOHKLKAKDQKPEVGEYAKLEKINAEQQLKIQELQEKLEK 180  
 QY 648 A-----AKERAERELEKLNQEDSSEGIKKLVAEERHLENKVKRL 691  
 DB 181 AVKASTEATELLQNIQAKERAERELEKLNQEDSSEGIKKLVAEERHLENKVKRL 240  
 QY 692 ETMERRENRLKDIQTKSQIQOMADKILEBEKREAOVSQAHLVHLKQEQHYEKEI 751  
 DB 241 ETMERRENRLKDIQTKSQIQOMADKILEBEKREAOVSQAHLVHLKQEQHYEKEI 300  
 QY 752 KYLDNQIKDLADKETLENMQHREBEAHEKGLISEQKAMNAMDSTKRSLEQRIVELS 811  
 DB 301 KYLDNQIKDLADKETLENMQHREBEAHEKGLISEQKAMNAMDSTKRSLEQRIVELS 360  
 QY 812 EANKLAANSLFTQRMNKAQEMISLRQOKFYLETQAGKLEAQNKLFEQLEKISHQDH 871  
 DB 361 EANKLAANSLFTQRMNKAQEMISLRQOKFYLETQAGKLEAQNKLFEQLEKISHQDH 420  
 QY 872 SDKNRLLETRLRVSLSEHEQKLEKQLTELQSLQERSQTLQAARAALESQLR 931  
 DB 421 SDKNRLLETRLRVSLSEHEQKLEKQLTELQSLQERSQTLQAARAALESQLR 480  
 QY 932 QAKTELEETAEAEIEIQTALTAHDEIQKFDALRNSCTVITDLBQLNQLTDEDAELNN 991  
 DB 481 QAKTELEETAEAEIEIQTALTAHDEIQKFDALRNSCTVITDLBQLNQLTDEDAELNN 540  
 QY 992 QNFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMELKTKTCTMLEEQ 1051  
 DB 541 QNFYLSKQLDEASGANDEIVQLRSEVDHLRREITEREMQLTSQKQTMELKTKTCTMLEEQ 600  
 QY 1052 VMDEALNDELLEKQWEAWRSVLGDEKSOQECRVREIQRMLDTEKQSRADQRIYES 1111  
 DB 601 VMDEALNDELLEKQWEAWRSVLGDEKSOQECRVREIQRMLDTEKQSRADQRIYES 660  
 QY 1112 ROVVAVLAVKHAETIALQALKEQKLEKESLSDKLNDELKHAMLENMARSLLQKLETE 1171  
 DB 661 ROVVAVLAVKHAETIALQALKEQKLEKESLSDKLNDELKHAMLENMARSLLQKLETE 720  
 QY 1172 RELKQRLLEEQAKLOQOMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYOLENIQVLY 1231  
 DB 721 RELKQRLLEEQAKLOQOMDLQKNHIFRLTQGLQEALDRADLLKTERSDLEYOLENIQVLY 780  
 QY 1232 SHEKVKMEGTISQOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLAL 1291

DB 781 SHEKVKMEGTISQOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLAL 840  
 QY 1292 EKEKARCAREEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQQIAMSIVSPSEH 1351  
 DB 841 EKEKARCAREEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQQIAMSIVSPSEH 900  
 QY 1352 QPSAMSLAPSSRRKESSTPEPSRRLKERMHNIPIHRFNGLNWRATKCAVCLDTVHF 1411  
 DB 901 QPSAMSLAPSSRRKESSTPEPSRRLKERMHNIPIHRFNGLNWRATKCAVCLDTVHF 960  
 QY 1412 GROASKLEQOVNCHPKCSTCLPATCGLPAEYATHFEAFCDKNNSPGLQTKPESSSLH 1471  
 DB 961 GROASKLEQOVNCHPKCSTCLPATCGLPAEYATHFEAFCDKNNSPGLQTKPESSSLH 1020  
 QY 1472 LEGMMKVPNNKGGQGWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGVSIIH 1531  
 DB 1021 LEGMMKVPNNKGGQGWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGVSIIH 1080  
 QY 1532 GAYGASELANAKA----- 1545  
 DB 1081 GAYGASELANAKADVPYILKMESHPTTCWPORTIYLLAPSPDKQWVTALESVVAGG 1140  
 QY 1546 -----EKAADA-----KLGNSLLKLEGGDRDLDMNCTLPFSDQVVLVGTE 1586  
 DB 1141 RVSREKAADAADWDCTSCERLPVWVEKLLGNSLLKLEGGDRDLDMNCTLPFSDQVVLVGTE 1200  
 QY 1587 EGYALNVNLNSLTHVPGIGAVFQIYIIKDLKLLMAGEERALCLVDVKKVKQSLAQSH 1646  
 DB 1201 EGYALNVNLNSLTHVPGIGAVFQIYIIKDLKLLMAGEERALCLVDVKKVKQSLAQSH 1260  
 QY 1647 LPAQDPISPIFAVKGCHLFGAGKIENGLCI CAAMP SKVILRYNENLSKYCIRKEIET 1706  
 DB 1261 LPAQDPVSPNIFAVKGCHLFAAGKIENSLCI CAAMP SKVILRYNENLSKYCIRKEIET 1320  
 QY 1707 SEPSCCHFTNYSLIGTNKFYIDMKQYTLBFLDKNDHSLAPAVFAASSNSFPYSIVQ 1766  
 DB 1321 SEPSCCHFTNYSLIGTNKFYIDMKQYTLBFLDKNDHSLAPAVFAASSNSFPYSIVQ 1380  
 QY 1767 VNSAQOREEYLLCFHEFGVFDVSYGRRSRDDELKWSRLPLAFAYREPFLVTHFNSLEVI 1826  
 DB 1381 ANSTGQOREEYLLCFHEFGVFDVSYGRRSRDDELKWSRLPLAFAYREPFLVTHFNSLEVI 1440  
 QY 1827 EIQRSSAGTPARAYLDIPNPRYLGAISSGATYLLASSYQDKLRVCCCKNLVKGSTEH 1886  
 DB 1441 EIQRSSAGTPARAYLEIPNPRYLGAISSGATYLLASSYQDKLRVCCCKNLVKGSTEH 1500  
 QY 1887 HRGPSTSRSPNKGPPPTNYEHITKRVASSPAPPEGSPHREPSTPHRY--REGRTELRR 1944  
 DB 1501 HRVPTSR--SPNKGPPPTNYEHITKRVASSPAPPEGSPHREPSTPHRYDREGRTELRR 1559  
 QY 1945 DKSPGPLEREKSPGRMLSTRERSPGRLEFDSRGLPAGAVRTPLSQVNVKRGQSA 2002  
 DB 1560 DKSPGPLEREKSPGRMLSTRERSPGRLEFDSRGLPAGAVRTPLSQVNVKWDQSS 1617

RESULT 5

Q8C1J1 ID Q8C1J1 PRELIMINARY; PRT; 1032 AA.  
 AC Q8C1J1;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to citron (Hypothetical protein).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID:10090;  
 RN [1]  
 RP SEQUENCE FROM N. A.  
 RC STRAIN=FVB/N; TISSUE=Breast tumor;  
 RA Strausberg R.;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC023775; AAH23775.1; -;

DR EMBL; BC051165; AAH51165.1; -.  
DR GO: 0005524; F:ATP binding; IEA.  
DR GO: 0005083; P:small GTPase regulatory/interacting protein. . .; IEA.  
DR GO: 0007242; P:intracellular signaling cascade; IEA.  
DR InterPro; IPR001180; Citron.  
DR InterPro; IPR005479; CPhase\_L\_D2.  
DR InterPro; IPR002219; DAG\_PE\_Bind.  
DR InterPro; IPR001849; PH.  
DR Pfam; PF00780; CNH; 1.  
DR Pfam; PF00130; DAG\_PE-bind; 1.  
DR Pfam; PF00169; PH; 1.  
DR SMART; SM00109; CL; 1.  
DR SMART; SM00036; CNH; 1.  
DR SMART; SM00233; PH; 1.  
DR PROSITE; PS00867; CPASE\_2; 1.  
DR PROSITE; PS00479; DAG\_PE\_BIND\_DOM\_1; 1.  
DR PROSITE; PS00881; DAG\_PE\_BIND\_DOM\_2; 1.  
DR PROSITE; PS00003; PH\_DOMAIN; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 1032 AA; 116919 MW; 59AEDCF659DA33B CRC64;  
  
Query Match 45.9%; Score 4819; DB 11; Length 1032;  
Best Local Similarity 91.5%; Pred. No. 1.2e-194;  
Matches 944; Conservative 10; Mismatches 10; Indels 68; Gaps 4;  
  
QY 1038 MEALKTCTMLEEQVMDLEALDELLEKERQWEAWRSVLGDEKSFQECRVRELQMLDTE 1097  
DB 1 MEALKTCTMLEEQVMDLEALDELLEKERQWEAWRSVLGDEKSFQECRVRELQMLDTE 60  
  
QY 1098 KOSRARADQRIETESQVVLAYKEHKAETLALQQAALKEQKLAESLSKINDLEKKHML 1157  
DB 61 KOSRARADQRIETESQVVLAYKEHKAETLALQQAALKEQKLAESLSKINDLEKKHML 120  
  
QY 1158 ENNARSLOQKLETERELKORLLEEQAKLOQMDLQKNHIFRITQGLQALRADLLKTER 1217  
DB 121 ENNARSLOQKLETERELKORLLEEQAKLOQMDLQKNHIFRITQGLQALRADLLKTER 180  
  
QY 1218 SDLEVOLENIQVLYGHEKVMGTTISQQTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALP 1277  
DB 181 SDLEVOLENIQVLYGHEKVMGTTISQQTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALP 240  
  
QY 1278 TVPQVLYNELKALEKARCAELAEALOKTRIELRSAREEAHAKATDHPSPATAPAR 1337  
DB 241 TVPQVLYNELKALEKARCAELAEALOKTRIELRSAREEAHAKATDHPSPATAPAR 300  
  
QY 1338 QIAMSALVRSPEHQPSSAMSLAPPSSRRKESSTPEEFSRRLKERMHNIPIHFNVGLNM 1397  
DB 301 QIAMSALVRSPEHQPSSAMSLAPPSSRRKESSTPEEFSRRLKERMHNIPIHFNVGLNM 360  
  
QY 1398 RATKCAVCLDTVHFGQASKCLEQVMCHPKCSTCLPATCGLPABYATHFTAPCRDKWN 1457  
DB 361 RATKCAVCLDTVHFGQASKCLEQVMCHPKCSTCLPATCGLPABYATHFTAPCRDKWN 420  
  
QY 1458 SPGLQTKPESSSLHLEGWVKVPRNNKRGQGWDRKYIVLEGSKVLIYDNEAREAGQRPVE 1517  
DB 421 SPGLQTKPESSSLHLEGWVKVPRNNKRGQGWDRKYIVLEGSKVLIYDNEAREAGQRPVE 480  
  
QY 1518 EFELCLPDGDVSIHGAVGASELANAKA----- 1545  
DB 481 EFELCLPDGDVSIHGAVGASELANAKA----- 540  
  
QY 1546 -----EKARADA----- 1572  
DB 541 QRVVTALESVAVAGRVSRKADARADHTSSEHQPVWVEKLLGNSLLKLEGDDRLDWNK 600  
  
QY 1573 TLPFSDQVVLVGTBEGLYALNVLKNSLTHVPVGIGAVFYIYIKLEKLLMIAGEERALCL 1632  
DB 601 TLPFSDQVVLVGTBEGLYALNVLKNSLTHIPGIGAVFYIYIKLEKLLMIAGEERALCL 660  
  
QY 1633 VDVKKVKQSLAQSHLPAQPDIPNITFEAVKGLFGAGKIENGLICIAAMPSKVILRYN 1692  
DB 661 VDVKKVKQSLAQSHLPAQPDIPNITFEAVKGLFGAGKIENGLICIAAMPSKVILRYN 720

QY 1693 ENLSKYCIRKEIETSEPCSCIHTFTNYSILICTNKFYEDIMKQYTLERFLDKNDHSLAPV 1752  
DB 721 DNLSKYCIRKEIETSEPCSCIHTFTNYSILICTNKFYEDIMKQYTLERFLDKNDHSLAPV 780  
  
QY 1753 FAASNSFPVSVIVQNSAGQREYYLLCFHEFGVFDVSYGRRSRTDDLLKWSRLPLAFAYRE 1812  
DB 781 FAASNSFPVSVIVQNSAGQREYYLLCFHEFGVFDVSYGRRSRTDDLLKWSRLPLAFAYRE 840  
  
QY 1813 PYLEVTHFNSLEVEIQARSSAGTPARAYLDIPNRYLGPALSSGAYLASSYQDKLRVI 1872  
DB 841 PYLEVTHFNSLEVEIQARSSAGTPARAYLDIPNRYLGPALSSGAYLASSYQDKLRVI 900  
  
QY 1873 CCKGNLVKESQTEHHRGPSTSRSPKRGPTTYNEHITKTVASSPAPPEGPSHPREPSTP 1932  
DB 901 CCKGNLVKESQTEHHRGPSTSRSPKRGPTTYNEHITKTVASSPAPPEGPSHPREPSTP 959  
  
QY 1933 HRY--REGRTLELRDKSPGRPLEREKSPGRMLSTRRSRSPGRLEFDSRGRPLPAGAVRTP 1990  
DB 960 HRYDREGRTLELRDKSPGRPLEREKSPGRMLSTRRSRSPGRLEFDSRGRPLPAGAVRTP 1019  
  
QY 1991 LSQVKNKRGGOSA 2002  
DB 1020 LSQVKNKWDQSS 1031  
  
RESULT 6  
O88937 PRELIMINARY; PRT; 494 AA.  
AC O88937;  
DT 01-NOV-1998 (TRENBERL. 08, Created)  
DT 01-NOV-1998 (TRENBERL. 08, last sequence update)  
DE 01-OCT-2003 (TRENBERL. 25, last annotation update)  
DE Rho/rac-interacting citron kinase short isoform.  
CIT OR CRIK-SK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99009084; PubMed=9792683;  
RA Di Cunto F., Calautti E., Hsiao J., Ong L., Topley G., Turco E.,  
RA Dotto G.P.;  
RT "Citron rho-interacting kinase, a novel tissue-specific Ser/Thr kinase  
encompassing the rho-rac-binding protein citron.";  
J. Biol. Chem. 273:29706-29711(1998).  
RC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AF086823; AAC72822.1; -.  
DR HSSP; P00517; 1STC.  
DR MGI; MGI:105313; Cit.  
DR GO: 0005524; F:ATP binding; IEA.  
DR GO: 0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO: 0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO: 0016740; F:transferase activity; IEA.  
DR GO: 0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000961; Kinase\_C.  
DR InterPro; IPR000715; Prot\_kinase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkinase.  
DR InterPro; IPR01245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00433; pkinase\_C; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00133; S\_TKC\_X; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 494 AA; 55742 MW; 2C60D2294B28A195 CRC64;  
  
Query Match 20.6%; Score 2165.5; DB 11; Length 494;

Best Local Similarity 87.8%; Pred. No. 1.6e-83;  
Matches 411; Conservative 26; Mismatches 30; Indels 1; Gaps 1;  
QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQFGKPPFMTQQQMSPLSRREGILDALFVLFE 60  
DB 1 MLKFKYGVNPPSEASAPIASRASRLNLFQFGKPPMTQQQMSALSREGMLDALFALFE 60  
QY 61 ECSQPALMKIKHVSNFVRKYSYDTIAELQELQPSAKOFVRSVLGCGHFAEVQVVRKATG 120  
DB 61 ECSQPALMKIKHVSFVKYSYDTIAELRELQPSARDFVRSVLGCGHFAEVQVVRKATG 120  
QY 121 DIYAMKMKKALLAQEOVSFFEEERNILSRSPWIPOLQYAFQDKNHLVLMEEYQPGG 180  
DB 121 DVYAMKMKKALLAQEOVSFFEEERNILSRSPWIPOLQYAFQDKNHLVLMEEYQPGG 180  
QY 181 DLLSLNRYEDQDENLIQYLAELILAVHSLVHMGYVHRDIKPNILVDRTGHILKLVDF 240  
DB 181 DFLSLNRYEDQDENLIQYLAELILAVHSLVHMGYVHRDIKPNILVDRTGHILKLVDF 240  
QY 241 GSAKAKNSKKNVAKLPIGTPDYMAPEVLTVMGDCKGTGVLCDWMSGVVAYEMVYGR 300  
DB 241 GSAKAKNSKKNVAKLPIGTPDYMAPEVLTVMGDCKGTGVLCDWMSGVVAYEMVYGR 300  
QY 301 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFLDLIQSLGQKRLKFEGLCCHPFF 360  
DB 300 TPTEGTSARTFNNIMNFORFLKFPDDPKVSSDFLDLIQSLGQKRLKFEGLCCHPFF 359  
QY 361 SKIDWNIRNSPPFPVPTLKSDDTSNFDPEKNSWSSPQSLSPSGSGBELPFGVGS 420  
DB 360 ARTDWNIRNSPPFPVPTLKSDDTSNFDPEKNSWSSPQSLSPSGSGBELPFGVGS 419  
QY 421 YSKALGILGRSESVSGLDSPAKTSSMEKLIKSKELQSDQKCHKM 468  
DB 420 YSKALGILGRSESVSGLDSPAKVSSMEKLLIKSKELQSDQKCHKV 467

RESULT 7  
O88527 ID O88527 PRELIMINARY; PRT; 448 AA.  
AC O88527;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Citron-K kinase (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Madaule P., Eda M., Watanabe N., Fujisawa K., Matsuoka T., Bito H.,  
RA Ishizaki T., Narumiya S.;  
RT "Role of Citron kinase as a target of the small GTPase Rho in  
RT cytokinesis".  
RL Nature 01-0-0(1998).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AF070065; AAC27932.1; -.  
DR HSPB; P00517; 1YDR.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000961; Pkinase.C.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser Thr pkinase.  
DR InterPro; IPR008271; Ser Thr pkin AS.  
DR InterPro; IPR001245; Tyr pkinase.  
DR Pfam; PF000649; Pkinase; I.  
DR Pfam; PF00433; Pkinase.C; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKC; 1.

SMART: SM00133; S\_TK X; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00109; PROTEIN KINASE ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
FT NON\_TER 448 448  
SQ SEQUENCE 448 AA; 50650 MW; 548CA5AEAE4A6394 CRC64;  
Query Match 20.5%; Score 2150.5; DB 11; Length 448;  
Best Local Similarity 90.6%; Pred. No. 6.2e-83;  
Matches 407; Conservative 21; Mismatches 20; Indels 1; Gaps 1;  
QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQFGKPPFMTQQQMSPLSRREGILDALFVLFE 60  
DB 1 MLKFKYGVNPPSEASAPIASRASRLNLFQFGKPPMTQQQMSALSREGMLDALFVLFE 60  
QY 61 ECSQPALMKIKHVSNFVRKYSYDTIAELQELQPSAKOFVRSVLGCGHFAEVQVVRKATG 120  
DB 61 ECSQPALMKIKHVSFVKYSYDTIAELRELQPSVRDFVRSVLGCGHFAEVQVVRKATG 120  
QY 121 DIYAMKMKKALLAQEOVSFFEEERNILSRSPWIPOLQYAFQDKNHLVLMEEYQPGG 180  
DB 121 DVYAMKMKKALLAQEOVSFFEEERNILSRSPWIPOLQYAFQDKNHLVLMEEYQPGG 180  
QY 181 DLLSLNRYEDQDENLIQYLAELILAVHSLVHMGYVHRDIKPNILVDRTGHILKLVDF 240  
DB 181 DLLSLNRYEDQDENMIQYLAELILAVHSLVHMGYVHRDIKPNILVDRTGHILKLVDF 240  
QY 241 GSAKAKNSKKNVAKLPIGTPDYMAPEVLTVMGDCKGTGVLCDWMSGVVAYEMVYGR 300  
DB 241 GSAKAKNSKKNVAKLPIGTPDYMAPEVLTVMGDCKGTGVLCDWMSGVVAYEMVYGR 299  
QY 301 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFLDLIQSLGQKRLKFEGLCCHPFF 360  
DB 300 TPTEGTSARTFNNIMNFORFLKFPDDPKVSSDFLDLIQSLGQKRLKFEGLCCHPFF 359  
QY 361 SKIDWNIRNSPPFPVPTLKSDDTSNFDPEKNSWSSPQSLSPSGSGBELPFGVGS 420  
DB 360 ARTDWNIRNSPPFPVPTLKSDDTSNFDPEKNSWSSPQSLSPSGSGBELPFGVGS 419  
QY 421 YSKALGILGRSESVSGLDSPAKTSSMEK 449  
DB 420 YSKALGILGRSESVSGLDSPAKVSSMEK 448

RESULT 8  
Q9VTY8 ID Q9VTY8 PRELIMINARY; PRT; 1854 AA.  
AC Q9VTY8;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CG10522 protein.  
GN CG10522.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley.  
EX MEDLINE=2019606; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Beres P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster L., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hsiao C., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP  
RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Buesam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Earfan D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hsiao C., Howland T.J.,  
RA Ibegwam M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirekas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of *Drosophila melanogaster* genome.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RP  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kamniker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of *Drosophila melanogaster* genome.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RN SEQUENCE FROM N.A.  
RP  
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RN SEQUENCE FROM N.A.  
RP  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003541; AAF49907.2; -;  
DR FlyBase; FBgn0036295; CG10522.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . . ; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR001180; Citron.  
DR InterPro; IPR002219; DAG PE-bind.  
DR InterPro; IPR000961; Pkinase\_C.  
DR InterPro; IPR000719; Prot\_Kinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR008271; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00780; CNH; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_Kinase; 1.  
DR SMART; SM00109; Cl; 1.  
DR SMART; SM00036; CNH; 1.  
DR SMART; SMC0220; S\_TK; 1.  
DR SMART; SMC0133; S\_TK\_X; 1.  
DR SMART; SMC0219; TY\_KC; 1.  
DR PROSITE; PS00081; DAG PE BIND DOM 2; 1.  
DR PROSITE; PS0107; PROTEIN KINASE\_ATP; 1.  
DR PROSITE; PS0011; PROTEIN KINASE\_DOM; 1.  
DR PROSITE; PS0108; PROTEIN KINASE\_ST; 1.  
KW ATP-binding; Transferase.  
SQ SEQUENCE 1854 AA; 211324 MW; 93B5B69FB63DDE1D CRC64;  
Query Match 17.5%; Score 1839.5; DB 5; Length 1854;  
Best Local Similarity 27.4%; Pred.No. 3.6e-69;  
Matches 552; Conservative 399; Mismatches 670; Indels 397; Gaps 71;  
QY 18 EPIASRASRLNLFPOGKP-----PFMTQQQMSPLS-----REG 50  
DB 6 EPISVTRARLNLLGKAGVCAKPGASGSGIPASTRSIVPVSTTSAAVAEAI 65  
QY 51 ILDALFVLFECSPALMK-IKHSNVRKYSDTIALQLQPSAKDFVRSVLGCHPA 109  
DB 66 LLDAFCLLYNECDKDTLKGDRNIAEFVNFRIIEETRLRYNADDFLITIGQYFG 125  
QY 110 EVQVVRKATGDIYAMKMKKALLAGEQVSPFEERNILSRSTSPWIPOLQVAFODKNH 169  
DB 126 NVHUVVERQNDIYAMKKIKKSVTTSQ-----VKEDIMINSRWLNQVAFQNDN 181  
QY 170 LYLMEEQPGGDLSSLNRYEDQDENLIQFYLAELILAVHSLVHLYVHRDKPENILV 229  
DB 182 LYLVMYMPGGDLSSLMSRH-GPDEDLARFYLAELTVALHTLHMGYVHRDKPENILI 240  
QY 230 DRTGHILVDFGSAKSNKNVNAKLPIGTDPYMAPEVLTVNMGD--GKGTGLDCDWM 287  
DB 241 DRFGHIXLADFGNAALDRDGHVLSLSPVGTPTDYIAPELQTISTYKLSKSMHDDVSCDYW 300  
QY 288 SVGVIAEMYGRSPFAEGTSARTFNNIM-----NFQRELFKFPDDPKVSDFLDLIQSL 341  
DB 301 SMGIIGELICETTFPHEDNVHETYSKILSHCBESHLKELISFPADLKSVNRYNLI 360  
QY 342 LCGQKERLKEGLCCHFPFSKIDWNNTNRPFPVPTLKSDDDTSNFDE-----PEKN 394  
DB 361 VTNPSKLSYERIKNHPFFSEIPWGSIRSQVPPPIPTVRSDDTSNFDGIRHKTREQG 420  
QY 395 SWVSSPCQLSPSGFSGBELPFVGFYSYKALGILGRSESVVGLDSPAKTSMMEKLLIK 454  
DB 421 VAKSLTNNKSNDFSGKDLFFIGYSFVH-----MEKSAISATTDKLEQK 466  
QY 455 SKE-----LQDSQDKCHKMEQMTLH---RRVSEVAVLQKVELKASQTSRLLQDL 507  
DB 467 LKELLQKLKTRENEISMLKDLRAQQLSRKTDNKSQVADAKMEIK--KLQOIKEKTM 524  
QY 508 ATYTECSLSKRSLE-QARMEVSQEDDKALQLLHDIREQSKLQEIK---EQEYQAOVEE 563  
DB 525 E--LTTCKTQIKTLQSSAKIDENMWKKEATITOLLRLNKKYBEAKIAEQRVEKQAD 582  
QY 564 MRLMWNQLEEDLVARRRSDIYSELRESRLAABEFKFKATECOH---KLKAKDQCKPE 620  
DB 583 KK-----QELASTLOKLDAREL-----EFNAKPEECKHLSMKLQNYKDM--- 621  
QY 621 VGEVAKLEKINAEQOLKIQELQEKLEAKAEARELEKLNQR-EDSEGRKKLVEAE 679

Db 622 -----LQIJK-EQNLKSETHNEBQRQMAELYEQLTDLRKKVRSDQDNRTWMEIKE 674  
Qy 680 RHSHLEKNV-----KLETHMEREN-----RUKDIOYKSOQ 711  
Db 675 IRELDSESISSKSTQBAKNATERNIEILRNEETIASNHELHAEKVLETKLQLKENE 734  
Qy 712 IQWADKILELEKHEKRAQVSAQHLVHLKQEGHYE-----EKIKVDNQIKKLADK 765  
Db 735 TQVRAECHRLERELQLAECQCLAESSLATQVPEYAPGSLTBLNATEDQLRADL--- 791  
Qy 766 ETLENNMORHEEBAHEKGLISLQKAMINAMDSKIRSLRQIRIVELSEANKLAANSFLTQ 825  
Db 792 -----LAAKESENHOKGR-----ADQLTQVTKLEQMLERFNEQSLSPKTS 832  
Qy 826 RNKKAQEMISELRQOKFYLETQAGLEAONRKLEROLEKISHODHSDKN-----RL 878  
Db 833 HSRKQEG-----ETVGMLEQNEKLEBKLAAVREQMIVERQAARTANLSW 880  
Qy 879 ELERLEUV-----SLEHEEQL-----ELKQLTELQSLQERESQITALQ- 920  
Db 881 KVEKLEELSEKLLARMLTREDRIKKVQNASDEAQMMLKTSQETQRESRIEELQ 940  
Qy 921 -----ARAALESQLOAKTELETTAAEHEIQALTAHRDEIQKEDALRNSCTWIDLE 976  
Db 941 ELAAKRDVLKHEKROWEKABQERKCKSEIIEHLANVHLEQOE-----TEL 988  
Qy 977 EQLNQLTE--DNAELANNQFY--LSKQLEASGANDEIVQLRSEVDHLRREITEREMOLT 1032  
Db 989 QKLRQIQSRFDGVTLEQNTVIRELQERESKRKANDSCLVLQKEL-----KQIT 1037  
Qy 1033 SQQWTEALKTCTMLEEQWDLBALNDELLEKEROWEAWSVLGDEKQFCRCRVEIQR 1092  
Db 1038 ---DNFQRLYACISITDSQLETEVETLTKSEQERKSKQSQDLDLHKEKRENDQDLEK 1094  
Qy 1093 MLQTEQSRARADQ---ITESQVVELAVKEHKBAILALQALKEQKLAESLSKND 1149  
Db 1095 QLTIVSEKELAQRAVLASEIDELRLNKEQKLVLAQQQLVEQ-----TNALFA 1147  
Qy 1150 LEKKHAWLE-MNARSLOQKLETERELKORLLEBAQKLOQMDLQKNHIFRL---TQGLQE 1205  
Db 1148 TORRAELLDQONANYEAQTADSNRENVLSKENARILSELFHKKEVEGNLQAEIRGLES 1206  
Qy 1206 ALBRADLLKTERSDLEVQENIQVLSHEKVKMEGTISQOTKLIDPLQAKM-DQPAKKK 1264  
Db 1207 A--QAN-LHAEDISQDTLAEKEQFTVQDRDKSNATLQKHLIDYLQKVEDLSAKKK 1263  
Qy 1265 GL-----FSRKKEDPALPTQV--PLQYNELKALEKAKAELEALQKTR--IELRS 1314  
Db 1264 TLADKLFGSSHTNKENVSPNDVESSILYRALKEELKREQKNSLLKEQLAQLNGTATLRS 1323  
Qy 1315 AREEAHRKATDHPHSTATARQOITAMSAIVSRPEHOPSMSLLAPSSRRKESSTPEE 1374  
Db 1324 PRISAANGSDAP-----KORPVSIAALPRSPQKQOQP---LKRTTSQVELKTTAEK 1373  
Qy 1375 FSRLEKRMHNIHPRFNVL---NMRAATKCAVCLDTHVFGQAQKLEQCMVCHPKGST 1431  
Db 1374 PTKVTIENQAH---HRFELAQESKYDANCVCVEKAVVAGSPFWKCKEVDTHRKCRS 1430  
Qy 1432 CLPATGCLFAEYATHEAFCDKDKNSPGLQTKERSSLSHLLEGWKMVPRNNKRGQGWDR 1491  
Db 1431 NVQSHCGSTKTPAP-----SADDLSISQSVSLTLDLS-VDVAGGTTSGGE--- 1474  
Qy 1492 KYIVLGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAKAEKAEAD 1551  
Db 1475 -YI---GS--LYSSDGAEE-----DQARKB----- 1493  
Qy 1552 AKLGNLSLLKLEGGDLDDNMCTLPFSD-QVILVGTTEGLYALNVLKNSLTHVPGIGAVFQ 1610  
Db 1494 -----IEYNCAFEVAEQIILLGCNTGLYAVHLDQSRVLHISGLSVSC 1537  
Qy 1611 IYIIXDLKLLMI--AGEBREALCLVDVKKV---OSLAQSHLPAQPDLS-----PNIEA 1660  
Db 1538 MSICKRLAKAIWGTVGKELYQC--DYQLESRCQSSSCHXFPVLETSAILPFPANRTPS 1595

Qy 1661 VKGCHLFGAGKIENGL--CICAMPSPKVVIILRYNENLSKYCIRKEIETSPSCSIHFTNY 1718  
Db 1596 EKWLVLISDEAENALDSVAJAATSTRIVILKYLKLMFKPVRALDTATPTVTFIFTRH 1655  
Qy 1719 SILICTNFEYIDMKQYTLFEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGOREEYLL 1778  
Db 1656 SAIVSSDFYIDLDNYAAEEFVLSDKSM-----ESTAKCOPLTAVRIS-----RQEYLL 1706  
Qy 1779 CFHEGCVDSYGRSRTDDLKWSRLPLAFAYRPPYLFVTHFNSLEVIEI----- 1828  
Db 1707 CFAEYGVFVDFBGCRSREPYDLNMYAPTGFYVRDPFLFISHYQSVQIVRLHRSFKEMAS 1766  
Qy 1829 --QARSSAGTP--ARAYLIDIPRYLGPAL--SSGAIYL 1861  
Db 1767 GDNASESEFELQRYL-----PHYMSLTLANSQDVNL 1800

RESULT 9  
OS4874  
ID 054874 PRELIMINARY; PRT; 1732 AA.  
AC 054874;  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Myotonic dystrophy kinase-related Cdc42-binding kinase.  
GN MRCK.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=98078670; PubMed=9418861;  
RA Leung T., Chen X.Q., Tan L., Manser E., Lim L.;  
RT "Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a  
RT Cdc42 effector in promoting cytoskeletal reorganization.";  
RL Mol. Cell. Biol. 18:130-140(1998).  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
DR EMBL; AF021935; AAC02941.1; -.  
DR PIR; T14039; T14039.  
DR HSP; P28867; IFTQ.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.  
DR GO; GO:0005529; F:sugar binding; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR001180; Citron.  
DR InterPro; IPR002219; DAG\_PE-bind.  
DR InterPro; IPR001304; Lectin C.  
DR InterPro; IPR000095; PAKbox/Rho-binding.  
DR InterPro; IPR001849; PH.  
DR InterPro; IPR003961; Pkinase C.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR Pfam; PF00780; CNH; 1.  
DR Pfam; PF00130; DAG\_PE-bind; 1.  
DR Pfam; PF00786; PBD; 1.  
DR Pfam; PF00189; PH; 1.  
DR Pfam; PF00089; pkinase; 1.  
DR Pfam; PF00433; pkinase C; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00109; C1; 1.  
DR SMART; SM00036; CNH; 1.  
DR SMART; SM00285; PBD; 1.  
DR SMART; SM00233; PH; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00133; S\_TK\_X; 1.

DR PROSITE; P850108; CRIB; 1.  
DR PROSITE; P800615; C TYPE LECTIN 1; 1.  
DR PROSITE; P800479; DAG PE BIND DOM 1; 1.  
DR PROSITE; P850081; DAG PE BIND DOM 2; 1.  
DR PROSITE; P850003; PH DOMAIN; 1.  
DR PROSITE; P8500107; PROTEIN KINASE ATP; 1.  
DR PROSITE; P850011; PROTEIN KINASE DOM; 1.  
DR PROSITE; P8500108; PROTEIN KINASE ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 1732 AA; 197063 MW; 65F62F7872ACCD3B CRC64;  
  
Query Match 13.8%; Score 1450; DB 11; Length 1732;  
Best Local Similarity 24.6%; Pred. No. 7.8e-53;  
Matches 521; Conservative 319; Mismatches 648; Indels 634; Gaps 75;  
  
QY 36 PFMTQOMSPLEGRGILDALFVFECSQPALMKIKHVSFVKYKSDTAELOEQPSAK 95  
DB 18 PAQTNGQC--FSVETLDDILICLYDECNPSLRREKNTILEYLEWAKPTSKVKQKRLHRE 75  
QY 96 DFEVSLVCGCHFAEVQVVRKATGDIYAMKMKKALLAQOVSPFFSEERNILSRSTSP 155  
DB 76 DFEILLKVGRAFGEVAVVKLVADKVFAFKILNKWMLKRAETACFREERDVLVNGDSK 135  
QY 156 WIPQOYAFQDKNHLVMEYQPGDILLSLNRYEDQDENLIQVLAELLILAVHSHLM 215  
DB 136 WITTLHYAFQDDNNLYLVMDYTVGGDLTLLSKFEDRUPENARFYLAEMVIAIDSVHQL 195  
QY 216 GYVHRDIKPNILVDRTGHIKLVDFGSAKMSNMVNAKLPIGTPDYMAPEVLTVWNGD 275  
DB 196 HYVHRDIKPNILMDMNGHILRADFGCLKMEDGTQSSVAVGTPDISPEILQAMB-D 254  
QY 276 GGTGTGLDCDWSVGVIAYEMLYIGRSPRAGTSARTFNIMNFQRLFPDD-PKVSDF 334  
DB 255 GGRGTGPECDWMSLGVCMYEMLYGTTPFYAESLVETYGKIMNHERFOFPTQVTDVSENA 314  
QY 335 LDLIOLGCGKRLKFEGL--CCHPPRSKDMNIRNSPPFVPTLKSDDDTSNFDPEP 391  
DB 315 KDLIRRLICSRHRLGQNGIEFKHPFGSGIDWNIRNCEAPYIPEVSSPTDTSNFDVD 374  
QY 392 E---KNSVSSPQCQLSPGSGEELPFVFGYSKALGILGRS-----BSV 434  
DB 375 DBCLKNSETMPPP--THTAFSGHLPFVFGTYTSTCVLSDRSLRVTAGTSLDLVDV 431  
QY 435 VSGLSPAKTSMWEKLLKSKELQDSQDKHMQEMLRHRYSB---VEA-----484  
DB 432 QRTLNNLATEYERRI-----KRLQEKLELRKLQESTQTVOALQYSTV 477  
QY 485 ---VLSQKEVELKASGTORSLLEQDLATYITECSLSKRSLEQARMEVSQEDDKALQLLHD 541  
DB 478 DGFLTASKDLKLSKEEIEKLKQVA---EVNHLEQOLEBAN-SVRELDLDAFRQIKA 532  
QY 542 IREQSRKLOEIKEOYQAOQVEEMRLMMNQLEEDLYSARRRSDLYSESLAESLAAEFKR 601  
DB 533 FEKQIKTLQOERE-----ELNKELVQASERLKNQSKELKDA-----568  
QY 602 KATECHKLLKAKDGKPPGVYAKLEKINAEQQLKIOELOEKLEAKAKERAERELEKIQ 661  
DB 569 ---HCQKRL-----ANQEFMEINE-----RLTELHTQKQLASHVYRDKS-BEVD 608  
QY 662 NREDSSEGRKXKLVAEERHSLKNVKRLETMERENRLKDDIOTKSOIQOQMAKILE 721  
DB 609 LVMQKAESLRQELRRAERAKKELEVHTEALIAEASKDKRLREQSRHYSKLEN-----662  
QY 722 LSEKHREAQVSAQHLEVLKQKEQHYEKKIYVLNDQ-----LKKDLAKETLENMMQRHE 776  
DB 663 LE-----GLKQKQISYSPGICIEHQEITKLTDLKKSIF-----YE 701  
QY 777 BEAHEXGKILSEQKAMINAMSKSRISLEQRIIVELSEANKLAANSLSFTQRNMKAQEEMIS 836  
DB 702 EBISKREGIHA-----SEIKNLKKELHD-SEGOQLANKELMW-----738  
QY 837 ELRQOKFYLETQAGKLEAQRNKLBEQLEKISHQDSDDKNRLLELTRLEVSLEHEBQKL 896

DB 739 -----LKDLEK-----TR-RESQSEREFEN 759  
QY 897 ELKQLTELQLSLQERESQLTALQAFAALESQLRQAkteLEETTAEEBEIQALTAHRD 956  
DB 760 EFKQQ-----YEREKVL-----771  
QY 957 EIQRKFDALRNSCTVITDLEEQNLQNTEDNAELNNQNFFYLSKOLDEASGANDEIVQRSE 1016  
DB 772 -----LTEENKLT-----SE 782  
QY 1017 VDHLRREITEREMQLTSQKQTMELKTTCTWLEBOVMDLEALNDELLEKEROWEAERSVL 1076  
DB 783 LD-----KLTS---LYESLSLRNHLEEEVKDLA---DKKESVAHWEA-----819  
QY 1077 GDEKSQFEQVRRELQRLMDTEKQSRARADQRITESRQVVELAVKHEKAEITALAQALKEQ 1136  
DB 820 -----QITEIIQWVSDEKDAAG-----YLOALASMTETE 850  
QY 1137 KLKAEISLSDKINDLE---KHAMLENNAR-SLOOKLETERELKORLLEBOAKLQQQMDIQ 1192  
DB 851 ALRNSSLGTRATDMPKWRFPKIDMSARLELSALDAEIRAQAIQOEELNKVK-----904  
QY 1193 KWHIFRLTQGLQBALDRADLLKTERSDELYOLENIQVLSHEKVKMBGTTISQOTKLIDFL 1252  
DB 905 -----ASNITECKLKDSEKKNLELLSEIQLIKTEBLRSEKGVHRDSQHSFL 954  
QY 1253 QAKVDQPAKXKGLFSRRKEDPALPTQVPLQYNELKLALEKAKARCAELBEALOKTRIEL 1312  
DB 955 AFLNTP-----DQDQPE---968  
QY 1313 RSAREEAAHRKATDHPHS--TPATARQOIAVSAIVRSEHOPSAmsLLAPSSRRKESST 1371  
DB 969 RS-----PCTPAGKGRRIADAPL--PVHTPTL-----RKGCPA 1002  
QY 1372 PEEFSRLKERMHHNIPIHRFNVGLNMEATCAVCLD-TVHFGRAQSKLEQVWCHPKCS 1430  
DB 1003 SAGFPPEK-----THQFEVKSFTAPTCHQCTSLMVGLIRGCGCEVCGFSCHITCV 1055  
QY 1431 TCLPATCGLPAEYATHTFEAFCDKKNSPGLQKEP-----SSSLHLEGNWVKVPRNK 1483  
DB 1056 NKAFTTCVPPE-----QTKPLGIDPQKGVGTAYEGHVRIPK-P 1094  
QY 1484 RG-QQGWDRKVIIVLEGSKVLIYDNEAREAGORPV-----EPELCLPDGDVSIHG 1532  
DB 1095 AGVKGHQRALAVVCDFKFLFYDIAEGKASQSPSSVISQVIDMRDEEFSSVSLASDVHA 1154  
QY 1533 A-----VGASELANTAKAE-----ADAKL-----LGNLLKLEGGDRRL 1568  
DB 1155 SRKDIPICIFRVTSQSL--SAPSDKCSILMLADSETERSKWWGVLSLHKVLKKNKFRDRS 1212  
QY 1569 -----DMNCTLP-----PSDOVLVGTTEEGLYALNVLKNSLTHVPGIGAVFQIY 1612  
DB 1213 VYVPKEAYDSTLPDIKTTQAAAIIDHERVALGNEEGLVVHVTKDEIRVGDNKKIHIQIE 1272  
QY 1613 IIKDLEKLLMAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKCHLFGAGKI 1672  
DB 1273 LIPSDQLVAVISGNRHRVLPFPMsALDGRETFYKLAEE-----TKGCTIIAAGKV 1322  
QY 1673 ENGL--CICAMPSKVILRYNENLSKYCRKIEITSEPCSCIHFTNYS--ILIGTKFY 1728  
DB 1323 RHGALSCLCVAMKEQVLCYELFQSKTRHKFKIQV--PCNVQMAIFSEHLVCG-----F 1376  
QY 1729 EIDMKQVTL-----EFLDKNDHSLAPAFVAASSNSFPVSIQVNSAGOREEYLLCFHE 1782  
DB 1377 QSGFLRPLNAGEGSPCNMLHSDHTLA--FITHQPMDAICAVEISN---KEYLLCFSS 1429  
QY 1783 FGVVDSYSGRRSRDLDLKWRLPLAFAYREPYLIVTHFNLSLEVIEIQARSAGTAPAYL 1842  
DB 1430 IGIYTDCCQRRSRQOELMWPANPSSCCYNAPYLSIYSENAVIDFDVNSMEWIQT-----L 1484  
QY 1843 DIPNRYLGAISGAIYSSGYQDKLRVICCKGNLVKES-----GTEHHRGPSTSRSP 1897  
DB 1485 PLKKVR---PLNTEGSLNLLG--LETIRLIYFKMMAEGDELVVPETSDNSRKQWNRIN 1539





[illegible]

Qy	1053	MDLEALNDELLEKEROEWAWSVLGDEKQFCEVRBELQEMLDTEKQSRARADQRTESR	1111
Db	851	QDLAAKK---ESVAHWEA-----QIAEIIQWVSDEKD---ARGYLQALAK	890
Qy	1113	QVVELAVKSHKABILALQALKEQKLKAESLSDKLNDLEKKHAMLEMMAR-SLQOKLETE	1171
Db	891	MTEEL-----EALRSSLSGRRTL-DPLWKV-RRSQKLDWSARLELQSALEAE	935
Qy	1172	RELUKQRLLEBQAKLQO-QMDLQKHIFRUTQGLQEBALRADLILKTRSDLEYQLENIQVL	1230
Db	936	IRAKVLVQBELRKVKDANILTESK-----LKDSEAKNRELLEEMEIL	977
Qy	1231	YSHEKVKMGCTISQQT--KLIDPLQAKMDQPAKKKKGLFSRRKEDPALTPQVPLQYNELK	1288
Db	978	---KKKMEKEFADTGLKLPDF-----QDSIFE-----YENTAP	1008
Qy	1289	LALKEKARCABELEBALQKTRIELRSARBEAHRKATDHPHS-----TPATARQOIAM	1343
Db	1009	LAHD-----LTFRTSSASEQTOAPKPEASPSMSVAASEQOEDMA	1048
Qy	1344	AIVRSPEHOPFSAMSLIAPSSRRKESSTPEFSRRLKERMHNIPIRFNVGLNMRAIKCA	1403
Db	1049	---RPPQPSAVPL---PTQALALAGPKKA-----HQFSIKSPSSPTQCS	1089
Qy	1404	VLCD-TVHFGROASKLEBQVMCHPKCSTCLPATCGLPAEYATHFTFEACRDMXNSPGLQ	1462
Db	1090	HCTSLMVGILTRQYACEVCSFACHVSKDGAPOVCPIPE-----Q	1130
Qy	1463	TKBP-----SSSLHLEGMKVPNRNKGQOQWDRKIYVLEGSKVIYD-NEAREAGQR	1514
Db	1131	SKEPLGVDVQRGIGTAYKGHVKVPKFT-GVKKGQRAYAVCDCKFLYDLPEGKSTQPG	1189
Qy	1515	PVTEFELCJPDGVSVTHGAVGASELANATAKEKA---EADAKILG-----NSLLKLEGDD	1566
Db	1190	VIASQVLDRDDEFSV-SSVLASDVTHATRRDIPCFRVTASLLGAPSKTSSLIILITENE	1248
Qy	1567	RLDNMNCTLPFSQOVLVGTGEGLYAL---NLVKNLSUTH-----FGIGAVFOIYII	1614
Db	1249	-----NEKKWVGILEGLQSLHKNLRLNQVVHPVLEAYDSSLPLIKAILTAIAV	1298
Qy	1615	KLEKLLMTAGBERALCLVDV-----KKVKQ-SLA-----QSHLPAQ	1650
Db	1299	-DADRIAV--GUEEGLYVTEVTRDVTIVRAADCKKHQCIELAPREKIVILLCGRNHVHLY	1355
Qy	1651	P-----DISPNIFE---AVKGCHLFGAGKIENGLCTCAAMPKSVWILRYNENLSKYCI	1702
Db	1356	PWSSLDGAGSFDIKLPETKGCQLMATATILKRNSTGTLFVAVKRLILCYEIQRTKPFHRK	1415
Qy	1703	EIBTSPCS--CHFTNYNIIILCTNKFYIDMKQYILE-----EFLDKNDHSLAFAVA	1754
Db	1416	FNSIVAPGSVQCLAVLRDLCLVG---YPSGFCLLSIGDGOQLNLVNPNDPSLA---FL	1468
Qy	1755	ASNSNPFVSIVQNSAGQEEYLLCFFHEFGVFDVSYGRRSRITDDLKWSRLPLAFA----	1809
Db	1469	SQGSFDALCAVELES---EYLLCFFSHVGLYVDPOGRRARAQELMWPAAPVACSCSP	1524
Qy	1810	---YRPEYLIVTHFNSLEVIE---IQARSSAGT-----PAR-----AYLD	1843
Db	1525	VTYISYGVDFVDVTRTMWVQITGLARIIFPLNSEGTLNLINCPEPPLIIVFKSKFSAVLN	1584
Qy	1844	IPNP-----RYL-----GPAISSGAIYLAASYQ-----	1866
Db	1585	VPDTSNKKQMLRTRSKRFVPKVEBERLQORREMLRDPRLRSKMISNPTNFNVAHM	1644
Qy	1867	---DKLVRVICKGNLVKESGTEHHRGPTSRSSPNKRGPTYNE-----	1907
Db	1645	GPQDGMQVLMDLPLSAVPPSQEERPGAPNTLA---RQPPSRNKNYIWSPPSCGSEPSVT	1701
Qy	1908	-----HITKRVASSAPPGEPSHPREPSTPHR	1934
Db	1702	VPLRSVSDPDQEDKPDSDSTKH---STFSNNSNPSPGPPSPNSPHR	1745

143	Q	D	QDENHLVLMVDYVVGDDLLTLISKFDKLPEDMARFVIGEMVLDAIDSIHQHLYHVHRDIK	202
164	Q	D	QDKNHLVLMBEYQFGGDLISLNNRYDQDLDENILIQFYLAELIILAVHSHVLMGVYVHRDIK	222
224	Q	D	QENILVDRTHIKVLDFGSAKNNNSKNNAKLPICGIPDYMAPEVLTVWNGDGKGTGYLD	283
203	D	B	QDNVLLDVNGHIRIADFGSLKJMDNDGTQVSSVAVGTPDYISPEILOAME--DGMKGKYPE	261
284	Q	D	QDWSVGVIAYEMTYIGRSPFAEGTSARTFNIMNFQRFLEKPPDD--PKVSSDFDLIQILL	342
262	D	B	QDWSLGVCMVEMLYGETPFPVABSLVETYGKIMNHBERFQPSHVTDVSEAKOLIQLI	321
343	Q	D	QCGKERIKFEGEL---CHPFPFSKIDWNIRNSPPFPVPTLKSDDDOTNSNFDPE---	396
322	D	B	QCSERRLGQNGIEDFKHAFEGGLNWNIRNLEAPYIPDVSSPDSPTNSFVDVDDDLNRNTE	381
397	Q	D	QVSSPCCQLSP---SGFSGEELPFVGFYSKALGILGFSESVVSGLDSPAKTSSMEKKLLI	453
382	D	B	QI-----LPGSHTGFGSLHLPFGFTFT-----TESCF\$--DRGSLKSIMQ\$NLT	425
454	Q	D	QKSELQDSQDKCHKMEQMTLHRRVRSEVAVL\$QKEVELKAS\$TOR\$LL\$EQD\$LATYITE	513
426	D	B	QKEDVQ--ROLEHSLQWE--AYERRI-----	447
514	Q	D	QCSSLKRSLEQAREMVSEDDKALQLLHDIREQSEKLOEI--KEQYIQAQVBEMLMNNQLE	572
448	D	B	Q-----RELEQEKLE\$KLOEQ\$TQVQ\$SLHGSSRALNSNRDK\$IKLNEEIERLKNKIA	502
573	Q	D	QEDV\$ARRR\$DLY\$EL\$R\$EL\$AAE\$FKKATECOHKL\$AKAQ\$KPEVGYAKLEKINA	632
503	D	B	QD-----SNRLERQL--EDTVLRQ-----	524
633	Q	D	QEOLKIOELQEKLEKAAKERAERALEKLQNRDPSSEGIRKXLEVAEBRRH\$SL\$ENKVRLE	692
525	D	B	QTORL\$G\$LEQHRVVRQ\$KEE-----LHKQLV\$EASER-----	555
693	Q	D	QTMERENRLKDDICTK\$SQOQOMADKLTLEEXKHREAQV\$SAQHLEVLHVKQ\$EQHYEEKIK	752
556	D	B	QK\$QAKELKDAHQ\$KALQ\$EFS-----ELN\$RVAELRAQ\$KQV\$R\$OLRDK\$E\$E\$MEVATQ	610
753	Q	D	QVLNQIKQDLADK\$ETLENMQR\$E\$E\$E\$E\$H\$E\$K\$G\$IL\$SEQKAMINADMSKIR\$UEQ\$RIVEL\$E	812
611	D	B	QKVD-----AMRQEM\$RAE-----KURKELEA-----QLDDVA\$E\$ASK	642
813	Q	D	QNKLAANS\$LFTOR\$NKAQ\$EM\$SEL\$BOQFYLETQ\$AKLEAQN\$RKELEQ\$LEKISHO\$HS	872
643	D	B	QERKLREH\$ENFC-----KQ\$E\$SELEALKV---KQ\$G\$GAGATLEH\$Q\$Q\$E\$IK\$S	688
873	Q	D	QDKNRLLE\$RL\$REV\$LE\$HE\$EQ\$LEK\$R\$OL\$TELQ\$LSQ\$ERESOL\$TAQ\$ARA\$LESOLRQ	932
689	D	B	Q-----ELEKVV---L\$FYZE---ELV\$R\$RASHVLE\$V\$N\$K\$V\$K\$E\$V\$H\$E\$SHQ\$ALQ\$K\$ILM	735
933	Q	D	QKTELE\$TTA\$AB\$EETI\$QAL\$TAH\$DEI\$Q\$R\$K\$PDAL\$NS\$CTVITD\$LE\$Q\$NQL\$T\$D\$N\$AE\$LNQ	992
736	D	B	QKDKLEK\$K\$R\$ER\$N\$E\$M\$E\$A\$V\$G-----TI\$KDKYERERAM\$LF\$DEN\$K\$K\$LTAE	779
993	Q	D	QNFYLS\$QOL\$D\$AS\$G\$AN\$DEI\$V\$OL\$R\$E\$V\$D\$H\$LR\$E\$ITER\$EM\$OLT\$S\$Q\$KT\$M\$E\$AL\$XTT\$CTM\$LE\$BQV	1052
780	D	B	QNEKLCSFVD-----KLT\$Q\$N\$R-----QLE\$DEL	801
1053	Q	D	QMLEALNDELLEK\$ERQ\$E\$W\$A\$R\$S\$V\$L\$G\$D\$E\$K\$OF\$E\$C\$K\$R\$E\$V\$E\$Q\$R\$M\$LTE\$K\$O\$S\$R\$A\$D\$O\$R\$IT\$E\$R	1112
802	D	B	QDL\$AAK-----E\$S\$V\$H\$K\$E\$A-----QIAEIIQW\$D\$E\$K\$D--ARGYLOAL\$ASK	841
1113	Q	D	QVV\$ELAV\$K\$H\$K\$B\$IL\$ALQ\$AL\$K\$E\$Q\$K\$LEK\$S\$D\$K\$INDLEK\$K\$H\$AM\$LE\$N\$N\$R--SLOQ\$K\$U\$LE\$T\$E	1171
842	D	B	QMT\$EEL-----EALR\$SS\$IG\$S\$R\$TL--DPL\$W\$K--RR\$Q\$K\$D\$M\$S\$AR\$LE\$Q\$S\$ALE\$E	886
1172	Q	D	QELKOR\$LL\$E\$E\$Q\$AL\$QO--QMDLQ\$K\$H\$IFR\$LTQ\$G\$EAL\$D\$RAD\$LL\$K\$T\$E\$R\$S\$D\$LE\$Y\$O\$LE\$N\$IQVL	1230
887	D	B	QIRAKQV\$Q\$EL\$K\$R\$Q\$AN\$T\$LE\$K-----J\$K\$E\$AK\$N\$R\$E\$LL\$E\$E\$M\$E\$IL	928
1231	Q	D	QY\$E\$K\$V\$K\$M\$E\$G\$T\$S\$Q\$Q\$T--KLIDFLQ\$K\$MDQ\$K\$K\$K\$GL\$F\$S\$R\$R\$E\$D\$ED\$PAL\$Q\$V\$P\$Q\$V\$N\$ELK	1288

```

Db 929  ---KKMEBEKFRADTGLKLPDF-----QDSIFE-----VFNTAP 959
Qy 1289 LALEKEKARCAEAEALQKTRIELRSARBEAAHRKATDHPHPS-----TPATARQOQIAMS 1343
Db 960 LAHD-----LIFRTSSASEQETQAPKEASPSMSVAASEQEDMA 999
Qy 1344 AIVRSEHQPSAMSLIAPSSRRKESSTPEERSRLKERHINIHRFNVLNMRATKA 1403
Db 1000 ---RPPQRESAVPL---PTQALVLGPKPKA-----HQFSIKFSFSTQCS 1040
Qy 1404 VCLD-TVHFGRQAKGLEQVMCHPKCSTCLPATCGLPAEYATHFTFAFCRDKMNSPGIQ 1462
Db 1041 HCTSLMWGLIRQYACEVCSFACHVCKGAPQVCPPE-----Q 1081
Qy 1463 TKEP-----SSSLHLEGMMKVPNNKRGQGWDRKYIVLEGSKVLIYD-NEAREAGOR 1514
Db 1082 SKRPLGVDVQRGIGTAYKGHVKPKPT-GVKKGWQRAYAVCECKFLYDLPEGKSTQPG 1140
Qy 1515 PVHEFELCLPDGVSIHGAVGASELANTAKA---EADAKLLG-----NSILKLEGDD 1566
Db 1141 VIASQVLDLDRDBFSV-SSVLASDVHATRRDIPCFIRVTASLLGAPSKYTSLLILTENE 1199
Qy 1567 RLDNMCTLPSSDQVLVGTBEGYAL---NVLKNSLTHV-----PGIGAVFQIYII 1614
Db 1200 ---NEKKWVGILEGQSILHKURLNQQVHVPLEAYDSSSLPIKAILTAAIV 1249
Qy 1615 KDLKLLMIGERALLVDV-----KVKQ-SLA-----QSHLPAQ 1650
Db 1250 -DADTAV--GLBEGLYIVETRDVIRVADCKKHQIHELAPREKIVILLCGRNHHVHLY 1306
Qy 1651 P-----DISPNIPE-----AVKGCHLFGAGKIENGLCTCAAMPKSVILRYNENLSKYCIK 1702
Db 1307 PWSLDCAGSFDIKLPETKGCQMATATLKRNSGTCLFVAVKRLILCYEIQRTKPFHRK 1366
Qy 1703 EITSEPCS---CHFTNYSILITGNKFEIDMKQYILE-----EFLDKNDHSLAPAVA 1754
Db 1367 FNBIVAPGSVQCLAVLRDLVCG---YPSGGFLLSIQDGGQPLNVAENPDSLA---FL 1419
Qy 1755 ASSNSRPVSVQNSAGOREYLLCFHEFGVFDVSYGRSRTDCLKWSRLPLAFA----- 1809
Db 1420 SQSFDALCAVELES---EYLLCFSHMGLYVDPQGRARAGELMWPAAPVACSCSPH 1475
Qy 1810 ---YREPLYFVTHFNSLEVIE-----IQARSSAGT-----PAR-----AYLD 1843
Db 1476 VTVYSEYGVDFVDTMVEWQTIGLRIRPLNSEGTNLNLCBPPLIYFKSKFGAVLN 1535
Qy 1844 INP-----RYL-----GPAISSGAIYVASSYQ----- 1866
Db 1536 VPDTSNKSKQMLRTSKRNFVFKVPEERELQORREMLRDLPELRSKMINPINFNVHAM 1595
Qy 1867 ---DKLRVICCKGNLYKESGTEHRRGPTSSRSPNKRGPPTVNE----- 1907
Db 1596 GPGDGMQVLMPLSAVPPSQEERPGPATNLA---RQPPSRNKPYSWPSGGSPSVT 1652
Qy 1908 -----HITRVASSAPPAGPSHPREPSTPHR 1934
Db 1653 VPLRSMDDPDQDFKEPDSSTKH---STPSNSNPPSGPPSPNSPHR 1696

```

RESULT 13

```

Q7TT49
ID Q7TT49 PRELIMINARY; PRT; 1713 AA.
AC Q7TT49;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CDC42-binding protein kinase beta.
GN CDC42BPB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA Huang C.Q., Wu S.L., Cheng Z.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RW EMBL; AY27590; AAF34403.1; -.
KW Kinase.
SQ SEQUENCE 1713 AA; 194886 MW; 80C999262C96DAA6 CRC64;

Query March 13.18; Score 1373; DB 11; Length 1713;
Best Local Similarity 24.78; Pred.No.1.3e-49; Indels 546; Gaps 67;
Matches 484; Conservative 306; Mismatches 625;

Qy 44 SPLREGILDALFVLFEECSQPALMKIKHVSFNVRKYSYDTIAELOELQPSAKDFEVRSLV 103
Db 23 SLSVETLLDVLVCLYTECSHSALRRDKYVAEFLEWAKPFTQLVKDQLHREDFEIKVI 82
Qy 104 GCGHFAEVQVVRKATGDIIVAMKVMKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYA 163
Db 83 GRGAFGEVAVVPMKNTERIIVAMKILNKWMLKRAETACFFREERDVLVNGDCQMITALHYA 142
Qy 164 FODKNHLYLMEEYOPGGDLLSLNRYEDQDENLIQYLAELILAVHSVHLMGYVHRDIK 223
Db 143 FQDNYLYLVMDVYVGGDLTLISKFDKLPEDWARYIGEMVLADSIIHQLHYVHRDIK 202
Qy 224 PENILVDRTHIKLVDFGSAKKNNSKNVNAKLPIDGTPDYMAPEVLTVMGDKGKGYGLD 283
Db 203 PDNVLLDVNGHIRLADFGSCLKNDDGTQVSSVAVGTPDYISPEILQAME-DGMGKYGPE 261
Qy 284 CDWMSGVIAVEMYIGRSPRAEGTSARTENNINMFORFLKFPDD-PKVSSDFDLDTQSLL 342
Db 262 CDWWSLGVCMFEMLYGTGTPFYASLVETYKIMHNERFPFPHVTVSEAKDLQRLI 321
Qy 343 CGQKERLKFEG---CCHPFFSKIDMNNIRNSPPPVFTLTKSDDDTNSFDEPE---KNSW 396
Db 322 CSRERRLGQNGIEDFKKHAFFEGNLNENIRLEAPYIPDVSSPDSPTNSFDDVDLNRNIE 381
Qy 397 VSSSPCQLSP---SGFSGEELPFVGFYSKALGTLGRSESVVSGLDSPAKTSMEKLLI 453
Db 382 I-----LPPSGHTFGSLHLPFGTFTT-----TESCFS---DRGSUKSMIQSNTLT 425
Qy 454 KSKELQSDQKCHKWEQEMTLHRVRVSEVAIVLSQKEVELKASETQSRSLLEQDLATYITE 513
Db 426 KDDEVQDRLENSLQIEAVERRIE----- 449
Qy 514 CSSLKRSLQARMVVS---QEDDKALQLLHDIRE---QSRKLOEIKQEYQCAQVEEMELM 567
Db 450 -----LEQEKLESLKLOESTQTQVSLHGSTRALCNSNRDKKIK--RLNEELERMKSK 500
Qy 568 M---NOLK---EDLVARRRSDLYSELRESRLAAEFKKKATECQHKLLKAKDQGRPEV 621
Db 501 MADSNRLERQLEDVTLRQ----- 519
Qy 622 GEYAKLEKINAEQQLKTOELQEKLEKAAKRAERELKLNQREDSSGIRKLVAEERR 681
Db 520 -----EHEDSTQRLKGLKQYRLARQKSE-----LHKQLVEASER- 555
Qy 682 HSLKNVKKLETMERRENRLKDDIQTQSQOIQQWADKILEEKHREAQVSAQHLVHLK 741
Db 556 -----LKSQTKELKDAHQQRKALQEPS-----ELNEMALRQKQKVSQRLR 599
Qy 742 QKEQHYBEKIKVLDNQIKKOLADKETLENMMQREHEEAHEKGLKILSEQKAMINANDSKIR 801
Db 600 DKEEMEEMVAMQKIDS-MRQDIRKSE-----KSRKELEAR----- 632
Qy 802 SLEQRIVELSEANKLANSSILFTORNMKAQEMISELQKQFYLETQAGLEAQNKRLEE 861
Db 633 -LEDAVAEASKERKLREHSESFS-----KOMRELETUKV---KOGGRPGATLEHQ 681
Qy 862 QLEKISHQDSDKNRILLELETRLRREVSLHEHEQKLEKRLQTELQLSLQBPRESQLTALQA 921
Db 682 EISKIRS-----ELEKKV---LFYEE---ELVREASHVLEVKVNVKEVHESES 724
Qy 922 ARAALESOLRQAKTELEETTAABEETI-QALTAHRDRIQKFDALRNSCVITTDLESQLN 980

```

```
Db 725 HQALQKQVLMKDKLEKSKRHSRHESEAEIAGMKDKYER-----BRA 767
QY 981 QLTEDNAELNNQNFYLSKOLDASGAN----DEIVOLRSE---VDHLRREITEREMOLTS 1033
Db 768 MLFDENKKLTAEENKLCFSVDKLTQNRQLDELDQLASKESVAHWAEIAELIQWSD 827
QY 1034 QKQTMELKTTCTMLEEQWMDLEALNDELLEKERQWEAWRSVLGDEKSFQFCRVRELQRM 1093
Db 828 EKDARGYLQALASKMTE--ELETLRSSLSGRTLDPLWK-VRRSQKLDMSARL-ELQSA 882
QY 1094 LDTEKQSRARADQRTITESQVVELAVKEHKAETILAQALKEOKLKAESLSKLDNLEKK 1153
Db 883 LEAE-----IRAKQLVHEELRKVDKLTSLAFESKLESEAK-----917
QY 1154 HAMLENNARSLOOQUETERELKORLEEQAKLQOQMDLOKNHIPLTQLOGLADRADIL 1213
Db 918 -----NRELLEENQSLKRMEEK----FRADTGL-----942
QY 1214 KTERSLEYQLENIQVLYSHEKVMEGTISQOTKLDFLOAKMDQPAKKKGLFSRKED 1273
Db 943 -----KLPDF-----QD 949
QY 1274 PALPTQVPLQVNELKIALEKEKARCAELBEALQKTRIELRSAREBAHRKATDHPHPSTP 1333
Db 950 PIPE-----YFNTAPLAHDLTERTSASQDETQASKLDL-----SPSVSA 990
QY 1334 ATARQOIAMSAIVRSEPHQPSAMSLAPPSSRRKSSSTPEBSRRLKERMHHNIPIHRNV 1393
Db 991 TSTEQO---EDAARS-QORPSTVPL---PNTQALAMAGFKPA-----HQFSI 1031
QY 1394 GLNMRATKCAVCLD-TVHFGROASKLECOVWCHPKCSTCLPATCGLPAEYATHTEAPC 1452
Db 1032 KGFPSPTQSHCTSLMVLIRGYACEVCAFCHSVCKDSAPQVCPDPE-----1081
QY 1453 RDMNNSPGLQTKEP-----SSSLHLEGMVKVPRNNKRGQGWDRKYVILEGSKVLIYD 1505
Db 1082 -----QSKRPLGVDVQRGIGTAYGYVVKPKPT-GVKKGWQRAYAVVCDCKLFLYD 1131
QY 1506 -NEAREAGORPVEEELCLPCDGVSHGAVGASELANATAKAKA---EADAKLLG-----1556
Db 1132 LPEGSTQPGVTASQVLDLRDDEFV-SSVLASDVIHATRRDIPICFRVTASLGSPSKT 1190
QY 1557 NSLLKL---EGDDR-----LDMN-----CTLFPDQV-----1580
Db 1191 SLLILITENENKRWGVILEGLQALLHKNRLSQVHVHQAQYDSSLPLIKTVLAAIV 1250
QY 1581 ----VLVGTBGLYALNVLKNLTHVPGIGAVFYIYIKLEKLMIAGERALCLVDVK 1636
Db 1251 DGDRIAVGLEGLYVIELTRDIVRAADCKKYQIELAPKEKLIILLCCRNH-----1302
QY 1637 KVKQSLAQSHLPAQPDIGSPIEA-----VKGCHLFGAGKTENGLCICAAMPSKV 1687
Db 1303 -----HVHLYPWTSPDGAENFIDKLPETKGCQLIATGTURKSSSTCLFVAVKRL 1353
QY 1688 ILRYNENLSKYCIRKEIETSEPCS-----CIHF-TNYSILIGTNKIFYEIDMKQY 1735
Db 1354 VLYCETIQTTPHRRKFNEIVAPGHVQWVAMFKDRLCVGVPSPGFSLLSITQGDQPLDL---1410
QY 1736 TLEEFILKNDHSLAPAVFAASNSPFSVTQVNSAGQREYLLCHHEGVFVDSYGRSR 1795
Db 1411 -----VNPADPSLA---FLSQQSFDALCAVELKS-----EYLLCFSHWGLYVDPQGRSR 1458
QY 1796 TDDLKWSRLPLAFAYREPFLYFTHFNSLEVIQIARSSAGTPARAYLDIPNPRVYGPAS 1855
Db 1459 TQELAWPAAPVACSSSHVTVSEYGVDFVDVMTME-----WVQTIGLRIRPLNS 1510
QY 1856 SGAIVLASYQDKLRVICCKGNLVKESGTEHHRGFSSTRSS 1896
Db 1511 DGSLLNLCG--EPPLIIVFKN---KFSGTVLN-VPTDSDNS 1545
```

RESULT 14

```
054875
ID OS4875 PRELIMINARY; PRT; 1702 AA.
AC OS4875;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Myotonic dystrophy kinase-related Cdc42-binding kinase MRCK-beta.
GN MRCK-BETA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98078670; PubMed=9418861;
RA Leung T., Chen X.Q., Tan L., Manser E., Lim L.;
RT "Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a
RT Cdc42 effector in promoting cytoskeletal reorganization.";
RL Mol. Cell. Biol. 18:130-140(1998).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
DR EMBL; AF021936; AAC02942.1; -.
DR PIR; T14050; T14050.
DR HSP; P28867; 1PTO.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR002219; DAG-PE-bind.
DR InterPro; IPR000095; FAKBox/RhoBindg.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00130; DAG-PE-bind; 1.
DR Pfam; PF00786; PBD; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR ProDom; PD0000001; Prot_kinase; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00235; PBD; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00108; CRIB; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1702 AA; 194019 MW; C7AF80707563D319 CRC64;
```

```
Query Match 13.1%; Score 1372; DB 11; Length 1702;
Best Local Similarity 24.7%; Pred. No. 1.5e-49;
Matches 476; Conservative 298; Mismatches 609; Indels 546; Gaps 65;

QY 44 SPLSRGILDALFVLFECSQPALMKIKHVSNFVRKYSDTIALQLQPSAKDFEVRSLV 103
Db 23 SLSLVTLLDVLVCLYTECSHSLRRDKYVAEFLWAKPFTQLVQKMDQLHREDFEIKVI 82
QY 104 GCGHFAEVQVREKATGDIYAMKVMKKALLAQEQVFFEEERNILSRSTSPWIPQLQYA 163
```

Db 83 GRGAFGEVAVVMKNTERTIYAMKILNKWEMLKRAETACPREERDVLVNGDCQWITALHYA 142  
Qy 164 FQKXHLVMEVYQPGDILSLNRVEDOLDENILQFYLAELILAVSHVHMGVHRDILK 223  
Db 143 FQDENLYLVMDYVVGDDLLTLLSKEDKLPEDMARFYIGENVLAIDSLHQJHYHRDILK 202  
Qy 224 PENILVDRTHIKLVDPGSAKONSNKVMNAKLPICTPDYMAPEVLTVNMGDKGTGYLD 283  
Db 203 PDNVLLDVNGHIRLADFGSLCKWMDGTQSSVAVGTPDYISPEILQAME-DMGKYGPE 261  
Qy 284 CDWWSGVVIAYEMYIGRSFPAETSARTNNIMNFORFLKFPDD-PKVSDDFLDIQSL 342  
Db 262 CDWWSLGVCMYEMLYGETPFYAESLVETYGKIMNHEERFQFPHSHVDVSEEAADLIQRLI 321  
Qy 343 CGOKERLKPEGL---CCHPFFSKIDWNNIRNPPPPVPTLKSDDDTSNFDPE---KNSW 396  
Db 322 CSERRLQNGIEDFKHAFEGELANENIRNLEAPYIPDVSPSPTSINFDDVDDVLRNIE 381  
Qy 397 VSSSPCOLSP---SGFSGBELPVGFSYSKALGILGRSVSVGLSDSPAKTSMEKKLLI 453  
Db 382 I-----LPGSHGTGSLHLPPIGFTFT-----TESCFS--DRGSLKSMIQSNTLT 425  
Qy 454 KSELODSQDKCHKMQEOMTRLHRRYSVEAVLSQKEVELKASQTSRLLEQDLATYITE 513  
Db 426 KDEDVORDLNSIQIEAYERRIR-----LDMN-----CTLPSDQV--- 1580  
Qy 514 CSSLKRSLQARMEVS---QEDDKALQLLHDIRE---QSRKLOEIKQEQYQAVQEMRLM 567  
Db 450 -----LEQKLELKRKLQESQTQVSLHGTALGNSNRDKETK--RLNEBELERMSK 500  
Qy 568 M---NQLE---EDLVARRSDLYSELSRESLAAEFKPKATECOQKLLKAKQCKPEV 621  
Db 501 MADSNRLERQLEDTVTLRO-----LHKLQVLEASER- 555  
Qy 622 GEYAKLEKINAEQOLKIQELQEKLEAKAERARELEKLQNRDSDSEGIRKKLVAEERR 681  
Db 520 -----EHEDSTORLKGLEQYELARQKEE-----LHKLQVLEASER- 555  
Qy 682 HSLKVKRLNWERENRDKDIQTKSQIQOMADKILLEBKREAVQSAHLVHLK 741  
Db 556 -----LKSQTKELKDAHQKQALQFBS-----ELNERMAELRSQKQVSRQLR 599  
Qy 742 QKQHYEELIKVLDNOLKDLADKETLENMORHBEAHEKGIILSEQKAMINAMDSKIR 801  
Db 600 DKSENEVAMQKIDS-MRQDIRKSE-----KSKKELEAR----- 632  
Qy 802 SLQRIVELSEANKLAANSLFTQNNKAQEMISELROCKPYLETOAGKLEAQNKRLEE 861  
Db 633 -LEDVAEASERKRLREHSEFS-----KQERELETQV---KQGRGPGCARTLEHQ 681  
Qy 862 QLEKISHQDHSKRNLELETRLREVSLEBEQKLEK-----QLTELQSLQERESQLT 917  
Db 682 ETSKIRS-----ELEKKV-----LFYEBELVRERSHVLEKVNKVEHSEHQL 727  
Qy 918 ALQARAALLESQURAKTELETTAEABEIQALTAHRDEIQKFDALRNSCTVITDLEE 977  
Db 728 ALQKEVLMKLEKSK---RERHSEMEBAIGAM---KDKYER----- 764  
Qy 978 QLNQLTEDNALNNQNYLSKQDLDEAGAN-----DRIVOLRSE---VDHLRREITEREMQ 1030  
Db 765 ERAMLFDENKLLTAENEKLSFVDKLTQNRQLEBDQLAKGSVAHWEAQIAETIQW 824  
Qy 1031 LTSQKQTEALKTKTCMLBQVMDALNDELLEKEROEAWRSVLGDEKSFQECVRREL 1090  
Db 825 VSEKDKARGYLQALASQWTE---ELTLRSSLSGSRTLPLWK-VRRSQKLDMSARL-EL 879  
Qy 1091 QRMLDTEKQSRARADORITESQRVSLAVKEHKAETILALQOALKQKLAESLSKDLNDL 1150  
Db 880 QSALEAE-----IRAKOLVHEELRKVKDTSLAFESKLKESEAK----- 917  
Qy 1151 EKXHAMLENMARSLOQKLETERELKORLLEEQAKLOQMDLQKHIFRITQGLQALDRA 1210  
Db 918 -----NRELLEEQSLKKRVEEK-----FRADTGL----- 942

Qy 1211 DLLKTERSDEYOLENTQVLYSHEKVKMEGTISQOTKLIDFLOAKMDQPAKKKKGLFSRR 1270  
Db 943 -----KLPDF----- 947  
Qy 1271 KEDPALPTQVLPQVNEUKLALKEKAKCAELBALQKTRIELRSAREEAAHRAKATDHPHP 1330  
Db 948 -QDPIFE-----YFNAPLAHDLTFRTSSASDQETQASKDL- 987  
Qy 1331 STPATARQIAMSALVRSPEHQPSAMSLAPPSSRRKESSTPEEFRRRLKERMHNIPHR 1390  
Db 988 SVATSTEQ---EDAARS-QORSTVPL---PNTQALAMAGPKPKA-----HQ 1028  
Qy 1391 FNVGLNVRATKCAVCLD-TVHFRQASKCLCEQVMCHPKCSTCLPATCGLPAYATHFTE 1449  
Db 1029 FSIKSPSPITQCSHCTSLMVGLIRQGVACEVCAFSCHVSCDSAPQVCPIPPE----- 1081  
Qy 1450 AFCEDKNSPGLQTKEP-----SSSLHLEGWVKVPRNNKRGQGWDRKYIVLEGSKVL 1502  
Db 1082 -----QSKRPLGVVDVORGITAYKGVKVPKPT-GVKKGWORAYAVCDCKLF 1128  
Qy 1503 IYD-NEAREAGQRFVEEFELCLPDGVSIHGAVGAGELANTAKAEKA---EADAKLLG-- 1556  
Db 1129 LYDLPECKSTQPGVIASQVLDLRDDEFAV-SSVLASDVIHATRRDIPCIPTVTSASLLGSP 1187  
Qy 1557 ---NSLILKL---EGDDR-----LDMN-----CTLPSDQV--- 1580  
Db 1188 SKTSSLLITENENKRWVGLLEGLOAILHKRLRSQVHVHVAQEAVDSSPLIKTVLAA 1247  
Qy 1581 -----VLVTGEBGLYALNVLNLSLTHVPGIGAVFIYIKLEKLLMAGEERALCLV 1633  
Db 1248 AIVDGDRIANGLEBGLVIELTRDVIIVRAADCKVQVIELAPKEKILLCGRNH----- 1302  
Qy 1634 DVKKVKSQSLAQSHLPAQPDISPNIPEA-----VKGCHLFGAGKIENGLCICAAPS 1684  
Db 1303 -----HVHLYPWTSTFDGAEASNFDIKLPETKGCQLIATGTLRKSSTCLFVAV 1350  
Qy 1685 KWTILRYNENLSKYCIKKEITETSEPCS-----CIHF-TNYSILIGTNKFYEIDM 1732  
Db 1351 KRLVLCVEIGETQFPHKNEIIVAPGHVQWAMFKORLVCVGPSPGSLLSIQDGGPLDL 1410  
Qy 1733 KQYTLFEBLDKNDHSLAPAVFAASSNFPVSVIVQVNSAGOREBYLLCFHFEFGVVDYSGR 1792  
Db 1411 -----VNPADPSLA---FLSQSPDALCAVELKS---EYLLCFSHMGVYVDPQGR 1455  
Qy 1793 RSRTDDLKWSRLPLAFAYREFYLVTHFNSLVEIEIQARSSAGTPARAYLIDIPNRYLGP 1852  
Db 1456 RSRTQELMWAAPVACSSSHVTYVSEYGVDFVDTME-----WVQTIGLRIRP 1507  
Qy 1853 AISSGAIYL 1861  
Db 1508 LNSDGSINL 1516

## RESULT 15

Q7TT50 PRELIMINARY; PRT; 1713 AA.  
AC Q7TT50;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CDC42-binding protein kinase beta.  
GN CDC42BPB.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Huang C.Q., Wu S.L., Cheng Z.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY277589; AAF34402.1; -.  
KW Kinase.

[illegible]

Qy	1035	KQTWEAKTCTTMLBEOQVMDLEALNDELLEKERQWEANRSVLDGDEKSQFECRVRELOQML	1094
Dy	829	KBARGYLQALASKWTE---ELETRRSSSLGSRITLDPLWK---VRRSQKLDMSARL-ELQSA	883
Qy	1095	DTKQSPARADORITRESQVVELAVKEHKAELLALQCALKEQKKAESLSDKLDNLEKKH	1154
Dy	884	EAE-----ITAKQVQBELRKVKDSSJAFESKUKESAK	917
Qy	1155	AMLEMNARSLOQKLETERELKQRLBEEQAKLOQMDLOKNHIFRLTQGLQEAALDRADLLK	1214
Dy	918	-----NRELLEMQOSLRKRMEEK---FRADTGLK-----L	944
Qy	1215	TERSDLEYOLENTQVLYSHEKVKWEGTISQOTKLIDFLOAKDQDOPAKKKGLFSRKEDP	1274
Dy	945	PDFQDSIFEYNTAPLAHDLTFTTSSASQOETO-----ASKMDLSPSVVATSTEQOEM	999
Qy	1275	ALPTQVPOVYNELKLALKEKARCAELEBALQKTRIELRSARBEAAHRKATDHPHSTPA	1334
Dy	1000	ARQORP-----SPVLPFSTQA	1011
Qy	1335	TARQQTAMSAIVSRSPHQPSAMSLAPPASRRRKESSTPEFSRRLKERMHNIHPHFNVG	1394
Dy	1017	LA-----NAGPKPK-----AHQFSIK	1032
Qy	1395	LNMRATKCAVCLD-TVHGRQASKCLECQVMCHPKGKSTCLPATCGLPAEYATHFTFAFCR	1453
Dy	1033	SPFSPTQCSHCITSLMVGTLRQGVACEVCAFSCVSHVCKDSAPQVCPTPE	1081
Qy	1454	DKMNSPGLQKTP-----SSSLHLHEGMKVPNRNKRQGOQWDRKYIWLVEGSKVLIYD-	1505
Dy	1082	-----QSKRPLGVDVORGIGTAYKGVYKVPKPT-GVKKGQWQRAYAVVCDCKLFLYDL	1132
Qy	1506	NEAREAGQRPVEFEICLPDGDVSIHGAVGASLANATAKAERK---EADAKLIG-----N	1557
Dy	1133	PEKSTQPGVVASQVJDLRDEEPAV-SSVLASDVIIHATRRDPCIFRVTASLUGSPSKVS	1191
Qy	1558	SLLKLEBDRDLNMNCTLPFSDDQVVLVYGTBEGLYAL---NVLKNSLPHV-----PGI	1605
Dy	1192	SLLIITENE-----NEKRKWKYGILEGLQAILHKRLKNSQVVHVAAQAYDSSLPLI	1241
Qy	1606	GAVFOYIILKLEKLMIAGERALCLV-----DVKKV-----KSLA-----	1643
Dy	1242	KAVLAAAIIV-DGDIRIAV-GLBEGLVVIELTRDVIVRAADCKKVQYIELAPKEKIAILLC	1298
Qy	1644	--QSHLPAOPDISPNI-FEA-----VKGCHLFAGKIEINGLCICCAAMPKSVWILRYN	1692
Dy	1299	GRNHVHLVFWSSFDCAEASNFDIKLPEYKGCOLIATGILRKSSSTCLFVAVKCLILCYE	1358
Qy	1693	ENLSKYCIKREIETSPCS-----CHHF-TNYSILGTNKFYIDMKQYTLBEF	1740
Dy	1359	IQRTKPFHRKFGBELVAFGHVQWMAVFKDRCLCVGYPGSGFSLLSITQGDGPELDL-	1410
Qy	1741	LDKNDSHSLAPVFAASSNFPVSIYOVNSAGOREEYLLCFHFGVFPDVSYGRSRSRTDCLK	1800
Dy	1411	VNPTDPSLA---FLSQOSFDALCAVELKS-----EYLLCFSEHGLYVDPQGRHRSRQOELM	1463
Qy	1801	WSRLPLAFAYRPFYLVTHFNLSLEVIEIOARSAGCTPAPAYALDIPNRYLGPASSGAYI	1860
Dy	1464	WPAAPVACSCSPHVVYSEYGVDPDVRTME-----WVQTIGLLRIRPLNSDGSLN	1515
Qy	1861	LASSYQDKLRVICCKGNLVKESGTEHHRGCPSTRSPNKRGPPTYNEHITKRVASSPAPP	1920
Dy	1516	LIGCC--EPRLLIYFKN---KFGGTILN-VPDTSNKKQMLTRSKRRRVFKV-----	1562
Qy	1921	EGPSHPRPSTPHRYREGRTELR-----RDKSFGRLP-----REKSPGRMLSTRRE-RSP	1970
Dy	1563	-----PEE---ERLQORREMLRDELRSKVISNPTNFNVHVAHMGPDGMQVLMDLPLSA	1613
Qy	1971	GRLFEDSSSRGLRPAAGAVRPLSQ	1993
Dy	1614	APTQVEBKQGPAPLPRQPPGR	1636

Search completed: July 3, 2004, 10:17:33  
Job time : 113 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 20:02:52 ; Search time 14333 Seconds  
(without alignments)  
13696.642 Million cell updates/sec

Title: US-10-017-216-1  
Perfect score: 6574  
Sequence: 1 agagcgcagtgaggagat.....atcgagaatgtaggtttaga 6574

Scoring table: IDENTITY NUC  
Gapop 10\_0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:

- 1: em\_estba.\*
- 2: em\_estin.\*
- 3: em\_estim.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_htc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_htc.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pln.\*
- 20: em\_gss\_vrt.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_mam.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_phg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gss1.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	915.2	13.9	1011	13 BQ070955	BQ070955 AGENCOURT
2	815.6	12.4	956	13 BQ071141	BQ071141 AGENCOURT
3	810.2	12.3	830	13 BUI181633	BUI181633 AGENCOURT
4	750.6	11.4	849	12 B1253509	B1253509 602973370

5	716.4	10.9	920	13 BX342268	BX342268
6	692	10.5	879	13 BQ228524	BQ228524 AGENCOURT
7	688.8	10.5	958	11 EC031156	EC031156 Mus muscu
c 8	629.2	9.6	652	10 BF905370	BF905370 IL3-MT026
9	606.6	9.2	881	12 BG976452	BG976452 602846269
10	589.8	9.0	746	14 CA749290	CA749290 UI-M-FY0-
11	584.6	8.9	843	12 BG912161	BG912161 602812833
12	573.6	8.7	647	14 CF744580	CF744580 UI-M-GVO-
13	572.8	8.7	730	14 CF723360	CF723360 UI-M-GVO-
c 14	570	8.7	826	14 CD172337	CD172337 AGENCOURT
c 15	570	8.7	894	13 BQ421272	BQ421272 AGENCOURT
c 16	570	8.7	2426	11 BC018818	BC018818 Homo sapi
17	566.4	8.6	1085	12 EM904785	EM904785 AGENCOURT
18	557.8	8.5	1201	13 BX449104	BX449104 BX449104
19	556.8	8.5	640	13 BQ807302	BQ807302 NISC_kk01
20	555.6	8.5	963	13 BX437603	BX437603 BX437603
c 21	544.8	8.3	778	10 BE910475	BE910475 601503793
c 22	540	8.2	540	9 AA746301	AA746301 oa56C11.f
c 23	534	8.1	805	14 CF126292	CF126292 UI-HF-ETO
c 24	523	8.0	726	10 BE735426	BE735426 601304218
25	520.6	7.9	1044	9 AL532373	AL532373 AL532373
c 26	514	7.8	854	13 BUI195960	BUI195960 AGENCOURT
27	511.8	7.8	775	13 BUI105833	BUI105833 603005490
c 28	511.4	7.8	954	12 EG513123	EG513123 602811735
29	511	7.8	804	13 BU236617	BU236617 603411670
30	510.4	7.8	933	13 BU232508	BU232508 603408272
31	508.6	7.7	547	14 CF744826	CF744826 UI-M-GVO-
c 32	507	7.7	564	12 BM826567	BM826567 K-EST0098
33	506.6	7.7	965	9 AL580381	AL580381 AL580381
c 34	502	7.6	940	12 BG744767	BG744767 602722887
c 35	497.4	7.6	572	9 AA034099	AA034099 z106108.f
c 36	492	7.5	624	9 AA044766	AA044766 zk67d10.f
37	491.6	7.5	875	13 BX723780	BX723780 BX723780
38	490.8	7.5	875	13 BU900029	BU900029 AGENCOURT
c 39	488.8	7.4	1127	12 BM544893	BM544893 AGENCOURT
40	484.2	7.4	820	13 BQ941506	BQ941506 AGENCOURT
41	473	7.2	647	10 BE207065	BE207065 BE207065
42	472.6	7.2	657	13 BQ831488	BQ831488 LL61n2149
43	471.6	7.2	638	14 CD348416	CD348416 UI-M-FY0-
44	471.4	7.2	1001	9 AL579049	AL579049 AL579049
c 45	466.6	7.1	681	10 AW605350	AW605350 QV3-DT004

ALIGNMENTS

RESULT 1  
BQ070955  
LOCUS BQ070955 1011 bp mRNA linear EST 02-APR-2002  
DEFINITION AGENCOURT\_6855847 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:5923441  
S., mRNA sequence.

ACCESSION BQ070955  
VERSION BQ070955.1 GI:19900001

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1011)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-f@mail.nih.gov](mailto:cgapbs-f@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2091 row: p column: 02

High quality sequence stop: 634.

[illegible]



```
Db      601 TGTCCAAAACATCGATGAGGCTTCTGGCGCCAAACGACGAGATTGTACAACTGCGAAGTG 660
QY      3065 AAGTGACCATCTCCCGCGGAGATCACGAAACGAGAGATGAGCTTACCACCGAAGC 3124
Db      661 AAGTGACCATCTCCCGCGGAGATCACGAAACGAGAGATGAGCTTACCACCGAAGC 720
QY      3125 AAGCATGAGGCTCTGAAGACCACTGACCATGCTGAGGAAACAGGTCATGGATT-G 3183
Db      721 AAGCATGAGGCTCTGAAGACCACTGACCATGCTGAGGAAACAGGTCATGGATTGG 780
QY      3184 GAGGCCCTAAACGATGAGCTGTAGAAAAAGACGCGCAGTGGAGGCC 3231
Db      781 GAGGCCCTANACGATGAGCTGTAGAAAAAGACGCGCAGTGGAGGCC 828

RESULT 4
BI253509 849 bp mRNA linear EST 17-JUL-2001
LOCUS 602973370F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112737 5',
DEFINITION mRNA sequence.
ACCESSION BI253509
VERSION 1
KEYWORDS GI:14805003
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 849)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgs.nhl.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11273 row: 1 column: 18
High quality sequence start: 2
High quality sequence stop: 772.
Location/Qualifiers
1. 849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5112737"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN
Query Match 11.4%; Score 750.6; DB 12; Length 849;
Best Local Similarity 98.5%; Pred. No. 7.7e-169;
Matches 789; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY      4677 ACTGCTTGGAACTCCCTGCTGAACTGGAAGTGATGACCGTCTAGACATGAATGCGAC 4736
Db      3 ACCGCTCGGAAATCCCTGCTGAACTGGAAGTGATGACCGTCTAGACATGAATGCGAC 62
QY      4737 GCTGCCCTTCAGTGACCACTGGTGTGTGGTGGCCACCGAGGAGGCTCTACGCCCTGAA 4796
Db      63 GCTGCCCTTCAGTGACCACTGGTGTGTGGTGGCCACCGAGGAGGCTCTACGCCCTGAA 122
QY      4797 TGTCTTGAAAACTCCCTAACCATCTCCAGGAATGGAGCAGTCTTCCAAATTATAT 4856
```

```
Db      123 TGTCTTGAAAACTCCCTAACCATCTCCAGGAATGGAGCAGTCTTCCAAATTATAT 182
QY      4857 TATCAAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCGCACTGTGTCTTGT 4916
Db      183 TATCAAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCGCACTGTGTCTTGT 242
QY      4917 GGAGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACAT 4976
Db      243 GGAGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACAT 302
QY      4977 CTACCCCAACATTTTGAAGCTGTCAAGGCTGCCATTTTGGGGCAGGCAAGATTGA 5036
Db      303 CTACCCCAACATTTTGAAGCTGTCAAGGCTGCCATTTTGGGGCAGGCAAGATTGA 362
QY      5037 GAACGGGCTCTGTCATCTGTGAGCCATGCCAGCAAAAGTCTCATTTCCGCTTACAACGA 5096
Db      363 GAACGGGCTCTGTCATCTGTGAGCCATGCCAGCAAAAGTCTCATTTCCGCTTACAACGA 422
QY      5097 AAACCTCAGCAAAATCTGTCATCCGAAAGAGATAGACCTCAGAGCCCTGCAGCTGTAT 5156
Db      423 AAACCTCAGCAAAATCTGTCATCCGAAAGAGATAGACCTCAGAGCCCTGCAGCTGTAT 482
QY      5157 CCATTTCACCAATTAACAGTATCTCTTGGAAACCAATAAAATTTTACGAAATCGACATGAA 5216
Db      483 CCATTTCACCAATTAACAGTATCTCTTGGAAACCAATAAAATTTTACGAAATCGACATGAA 542
QY      5217 GCAGTACACGCTCGAGAAATTCCTGGATAAAGATGACCATTCCTGGCACCTGCTGTGTT 5276
Db      543 GCAGTACACGCTCGAGAAATTCCTGGATAAAGATGACCATTCCTGGCACCTGCTGTGTT 602
QY      5277 TGCCGCTCTTCCAAACAGCTTCCTGTCATCAATCTGTCAGGTGAACAGCCAGGCGAGCG 5336
Db      603 TGCCGCTCTTCCAAACAGCTTCCTGTCATCAATCTGTCAGGTGAACAGCCAGGCGAGCG 662
QY      5337 AGAGGAGTACTTGTGTTTCCACGAATTTGAGTGTTCGTGGATTCTTACGGAAGACG 5396
Db      663 AGAGGAGTACTTGTGTTTCCACGAATTTGAGTGTTCGTGGATTCTTACGGAAGACG 722
QY      5397 T-AGCCGACAGAGATCTCAAGTGGAGTCTTACCTTTGGCTTTGCTTACAGAGAAC 5455
Db      723 TAAGCCGACAGAGATCTCAAGTGGAGTCTTACCTTTGG-CTTTGGCTTACAGAGAAC 780
QY      5456 CTTATCTGTTTGTGACCCACT 5476
Db      781 CTTATCTGTTTGGAGCCATT 801

RESULT 5
BX342268 920 bp mRNA linear EST 02-MAY-2003
LOCUS BX342268
DEFINITION BX342268 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK011VB20 5-PRIME, mRNA sequence.
ACCESSION BX342268
VERSION BX342268.1 GI:30334095
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 920)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6533.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK011DA10QP1&cluster=6533.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
```

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK011DA10QPl.		http://image.llnl.gov	
FEATURES	source	1. .920	
		Location/Qualifiers	
ORIGIN	Query Match Best Local Similarity 10.9%; Score 716.4; DB 13; Length 920; Matches 746; Conservative 0; Mismatches 31; Indels 1; Gaps 1;	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CS0DK011YB20"	
		/cell_type="HELA CELLS COT 25-NORMALIZED"	
		/cell_line="HELA"	
		/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"	
		/note="1st strand cDNA was primed with a NotI-oligo(dT)	
		primer. Five prime end enriched, double-strand cDNA was	
		digested with Not I and cloned into the Not I and EcoR V	
FEATURES	source	sites of the pCMVSPORT 6 vector. Library was normalized."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
		Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	/tissue_type="melanotic melanoma"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
		Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
		Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
		Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/db_xref="taxon:9606"			
/clone="IMAGE:6061485"			
/tissue_type="melanotic melanoma"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH_MGC_72"			
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;			
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.			
Average insert size 2 kb. Library constructed by Life			
Technologies."			
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
		Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/db_xref="taxon:9606"			
/clone="IMAGE:6061485"			
/tissue_type="melanotic melanoma"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH_MGC_72"			
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;			
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.			
Average insert size 2 kb. Library constructed by Life			
Technologies."			
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
		Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/db_xref="taxon:9606"			
/clone="IMAGE:6061485"			
/tissue_type="melanotic melanoma"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH_MGC_72"			
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;			
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.			
Average insert size 2 kb. Library constructed by Life			
Technologies."			
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6061485"	
		/tissue_type="melanotic melanoma"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_72"	
		/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;	
ORIGIN	Query Match Best Local Similarity 10.5%; Score 692; DB 13; Length 879; Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.	
		Average insert size 2 kb. Library constructed by Life	
		Technologies."	
		1. .879	
		Location/Qualifiers	
		/organism="Homo sapiens"	

Db 361 ACCTGGCGCTTGCTGCTGATATGCCACACACTTCACCGAGGCGCTTCTGCCGTGACAAA 420  
Qy 4384 ATGAATCTCCCGAGGTCTCCAGACAAGAGGCCAGAGCAGAGCTTGCACTGGAAGGTGG 4443  
Db 421 ATGAATCTCCCGAGGTCTCCAGACAAGAGGCCAGAGCAGAGCTTGCACTGGAAGGTGG 480  
Qy 4444 ATGAAGTGGCCCGAGGATACAAACGAGGACACAGAGCTGGGACAGGAAGTACATTGTC 4503  
Db 481 ATGAAGTGGCCCGAGGATACAAACGAGGACACAGAGCTGGGACAGGAAGTACATTGTC 540  
Qy 4504 CTGGAGGATCAAAAGTCTCTATTATGACAAATGAAGCCAGAGAGCTGGACAGAGCGCG 4563  
Db 541 CTGGAGGATCAAAAGTCTCTATTATGACAAATGAAGCCAGAGAGCTGGACAGAGCGCG 600  
Qy 4564 GTGGAAGAAATTTGAGCTGTCTCCCTCCGAGCGGGATGTATCTATTATGTCGCTTGGT 4623  
Db 601 GTGGAAGAAATTTGAGCTGTCTCCCTCCGAGCGGGATGTATCTATTATGTCGCTTGGT 660  
Qy 4624 GCTTCCGAATCGCAATACAGCAAGCAGA 4655  
Db 661 GCTTCCGAATCGCAATACAGCAAGCAGA 692

RESULT 7  
BC031156  
LOCUS BC031156 958 bp mRNA linear HTC 04-MAR-2003  
DEFINITION Mus musculus, similar to citron, clone IMAGE:4976752, mRNA.  
ACCESSION BC031156  
VERSION BC031156.1 GI:21411076  
KEYWORDS HTC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (base 1 to 958)  
Strausberg.R.  
Direct Submission  
Submitted (03-JUN-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapsb-remail.nih.gov](mailto:cgapsb-remail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
cdna Library Preparation: Life Technologies, Inc.  
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mdcpaxil.stanford.edu](mailto:mdcpaxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 59 Row: J Column: 6  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis  
This clone has the following problem: retained intron.

FEATURES  
Location/Qualifiers  
1..958  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4976752"  
/tissue\_type="mammary tumor, C3(1)-Tag model. Infiltrating  
ductal carcinoma, 5 month old virgin mouse."  
/clone\_lib="NCI CGAP Mam6"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

ORIGIN  
Query Match 10.5%; Score 688.8; DB 11; Length 958;  
Best Local Similarity 85.4%; Pred. No. 6.3e-154;  
Matches 781; Conservative 0; Mismatches 127; Indels 6; Gaps 1;  
Qy 5167 AATTACAGTATCCCTCATTGGAAACCAATAAATTCACGAATCGACATGAAGCAGTACACG 5226  
Db 40 AAGAAACGATCCCTCATTGGGACCAACAATTCATGAGATCGACATGAAGCAGTACACG 99  
Qy 5227 CTCGAGGAATTCCTGSGATAAGAATGACATTCCCTTGGCACCTGCTGTGTTTGGCGCTCT 5286  
Db 100 CTTGATGAGTTCCTGGCAAGAAGACGACATTCTCTTGGCACCTGCTGTGTTTGGCGCTCT 159  
Qy 5287 TCCAAACAGCTTCCTGCTCTCAATCGTCAGTGAAGCGAGCGGCGAGGAGGAGTAC 5346  
Db 160 TCCAAACAGCTTCCTGCTCTCAATGTCAGCGCAACAGCGCGGCGAGGAGGAGTAC 219  
Qy 5347 TTGCTGTGTTTCCACGAATTTGGAGTCTTCGTGATTTCTTACGGAAGACGTAGCCGCACA 5406  
Db 220 CTGCTGTGCTTCCACGAATTTGGGTCGTTCTGATTTCTTACGGAAGACGTAGCCGCACA 279  
Qy 5407 GACGATCTCAAGTGGAGTCTGCTTACCTTTGGGCTTTGCCCTACAGAGAACCTTATCTGTTT 5466  
Db 280 GATGATCTTAAGTGGAGTCTGCTTACCTCTGGGCTTGGCTTACAGAGAACCTTATCTGTTT 339  
Qy 5467 GTGACCCACTTTCAACTCAGTGAAGTAATGAGATCCAGGACGCTCTCTCAGCAGGAGCC 5526  
Db 340 GTGACTCACTTCAACTCCTTGGAGTCAATGAGATCCAGGCGCAGATCTCTCAGTGGGAGC 399  
Qy 5527 CCTCCCGAGCGTACCTGGACATCCCGAACCCCGCTTACCTGGGCGCTGCGATTCTCTCA 5586  
Db 400 CCTCCCGAGCATATCTGGAATTTCCAAACCTCGCTTACCTGGGCGCGGATTTCTCTCC 459  
Qy 5587 GGACGATTTACTTGGGCTCTCATATACAGGATAAATTAAGGTTCATTCTCTCAAGGGA 5646  
Db 460 GGACGATTTACTTGGGCTCTCATATACAGGACAAGTTAAGGTTCATATGCTCGAAGGA 519  
Qy 5647 AACCTCTGGAAGGAGTCCGGCAGTGAACACACCGGGGCGCTCCACCTCCCGCAGCAGC 5706  
Db 520 AACCTCTGGAAGGAGTCCGGCAGTGAACACACCGGGGCGCTCCACCTCCCGCAGCAGC 579  
Qy 5707 CCCAAAGCGAGCGCCACCCAGCATACAGAGCAGATACCAAGCGCGTGGCTCCAGC 5766  
Db 580 CCCAAAGCGAGCGCCACCAACATACAGAGCAGATACCAAGCGCGTGGCTCCAGC 639  
Qy 5767 CCAGCGCGCGCCGAAGCGCCAGCCACCCCGAGAGCCAAAGCAGACACCCACCGCTACC-- 5824  
Db 640 CCAGCGCGCACCGGAAGCGCCAGCCACCCCGAGAGCCAAAGCAGACACCGCAGCCGCTACCGA 699  
Qy 5825 ----GCAGGCGCGGACCGAGCTGCGCAGGGAAGAAGTCTCTCTGGCGCGCGCCCTTGGAGCGGA 5880  
Db 700 GACAGAGAGGGTCCGACAGAGCTGCGCAGGGAAGAAGTCTCCAGGCGCGCGCTCTGGAGCGG 759  
Qy 5881 GAGAAGTCCCGCGCGCGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGCGGTGTTT 5940  
Db 760 GAGAAGTCCCGCGCGCGATGCTCAGCAGCTAGAGAGAGCGGTCCCGCGGAGAGACTGTTT 819  
Qy 5941 GAAGACAGCAGCAGGCGCGGCTGCTCTGCGGAGCGGTGAGGACCGCGCTGTCTCCAGGTG 6000  
Db 820 GAAGACAGCAGCAGGCGCGGCTGCTCTGCGGAGCAGTGAAGGACCGCAGCTGTCTCCAGGT 879  
Qy 6001 AACAGGGAAGGCGCAGAGTCTCTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTAT 6060  
Db 880 AACAGGCTCTGGGACCGAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTAT 939  
Qy 6061 GACTGGAATAAAA 6074  
Db 940 AAAAAAAAAAAAAA 953

RESULT 8  
BF905370/c

LOCUS BF905370 652 bp mRNA linear EST 18-JAN-2001  
 DEFINITION IL3-MT0267-261200-410-H07 MT0267 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF905370  
 VERSION BF905370.1 GI:12296829  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 652)  
 AUTHORS Dias Neto, B., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-MT0267-261200-410-H07&t3=2000-12-26&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 624.  
 FEATURES  
 Location/Qualifiers  
 1..652  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="MT0267"  
 /note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ONESIES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
 ORIGIN  
 Query Match 9.6%; Score 629.2; DB 10; Length 652;  
 Best Local Similarity 98.0%; Pred. No. 1e-139;  
 Matches 637; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
 QY 2495 GGAACATGAAGCCCAAGAGAGATGATTCTGAACTCAGGCAACAGAAATTTACCTGG 2554  
 Db 652 GGAACATGAAGCCCAAGAGAGATGATTCTGAACTCAGGCAACAGAAATTTACCTGG 593  
 QY 2555 AGACACAGCTGGGAAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAGAAGA 2614  
 Db 592 AGACACAGCTGGGAAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAGAAGA 533  
 QY 2615 TCAGCCACCAAGACACAGTGCAGAGATCGCTCTGGAATTCGAGACAGATTCGGGG 2674  
 Db 532 TCAGCCACCAAGACACAGTGCAGAGATCGCTCTGGAATTCGAGACAGATTCGGGG 473  
 QY 2675 AGGTCACTAGACACAGGAGGAGGAGAACTCGAGCTCAAGGCCAGCTCAGAGAGTAC 2734  
 Db 472 AGGTCACTAGACACAGGAGGAGGAGAACTCGAGCTCAAGGCCAGCTCAGAGAGTAC 413  
 QY 2735 AGCTTCCCTGAGAGCGCGAGTCAAGTTGACAGCCCTGACAGCTCAGCGGCGGCC 2794

Db 412 AGCTTCCCTGACAGGCGCGAGTCAAGTTGACAGCCCTGACGGCTGCACGGCGGCGCC 353  
 QY 2795 TGGAGAGCAGCTTGGCCAGCGGAAGACAGAGCTGGGAAGAGACACAGCAGAGCTGAAG 2854  
 Db 352 TGGAGAGCAGCTTGGCCAGCGGAAGACAGAGCTGGGAAGAGACACAGCAGAGCTGAAG 293  
 QY 2855 AGGAGATCCAGGCACTCAGGCAATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTC 2914  
 Db 292 AGGAGATCCAGGCACTCAGGCAATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTC 233  
 QY 2915 GTAACAGCTGTACTGTATATCAGACAGCTGGAGGAGCAGCTAAACACGAGGAGACA 2974  
 Db 232 GTAACAGCTGTACTGTATATCAGACAGCTGGAGGAGCAGCTAAACACGAGGAGACA 173  
 QY 2975 ACGCTGAACCTCAACAAACAACTTCTACTTGTCTCAACAACTCGATGAGGCTTCTGCG 3034  
 Db 172 ACGCTGAACCTCAACAAACAACTTCTACTTGTCTCAACAACTCGATGAGGCTTCTGCG 113  
 QY 3035 CCAACGACGAGATTGTATACTCGAAGTGAAGTGGACCATCTCCGCGGAGATCAGG 3094  
 Db 112 CCAACGACGAGATTGTATACTCGAAGTGAAGTGGACCATCTCCGCGGAGATCAGG 53  
 QY 3095 AACGAGAGATGCAGCTTACGAGCCAGAGCAACAGGATGGAGGCTCTGAAG 3144  
 Db 52 AACGAGAGATGCAGCTTACGAGCCAGAGCAACAGGATGGAGGCTTGAACAAG 3  
 RESULT 9  
 LOCUS BG976452 881 bp mRNA linear EST 12-JUN-2001  
 DEFINITION 602846269P1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:4976752 5', mRNA sequence.  
 ACCESSION BG976452  
 VERSION BG976452.1 GI:14364089  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 881)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Jeffrey Green M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10970 row: j column: 17  
 High quality sequence stop: 841.  
 Location/Qualifiers  
 1..881  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4976752"  
 /sex="female, virgin"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Mam6"  
 /notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"  
 ORIGIN  
 Query Match 9.2%; Score 606.6; DB 12; Length 881;



Db 301 CAGGTGGTGGAGTGGCGGTGAAGAACACAAAGCGCGAGATTCTTGCTCTGCAGCAGCT 360  
QY 3415 CTCAAAGACGAGAAGCTGAAGCCGAGAGCCCTCTCTGACAACTCAATGACCTGAGAG 3474  
Db 361 CTCAGGACGAGAGCTCAAGCCGAGAGCCTGTGCGACAGCTCAACGACCTGAGAG 420  
QY 3475 AAGCATGCTATGCTTGAATGAATGCCGAACTTACAGCAGAACTGAGACTGAAACGA 3534  
Db 421 AAGCAGCCATGCTGAGATGAAGCCCGGAGCTTACAGCAGAACTAGACAGAGCGG 480  
QY 3535 GAGCTCAACAGAGGCTCTGAGAGCAGCAGCCAAATACACGACAGATGACCTGAG 3594  
Db 481 GAGCTCAACAGAGGCTCTGAGAGCAGCAGCCAAATACACGACAGATGACCTGAG 540  
QY 3595 AAAAAATCACTTTCCGCTGACTCAAGGACTGCAAGAGCTCTAGATGGGCTGATCTA 3654  
Db 541 AAGAACCATCTTTCAGACTGACGCAAGGCTGACGAGCGCTGACCGGCGATCTG 600  
QY 3655 CTGAGACAGAGAAGTCACTTGAGTATCAGCTGAGTGGAAACATTCAGGTTCTTATCT 3714  
Db 601 CTGAGACAGAGAAGGAGCAGCTGAGTACCAGCTGGAAGCAATTCAGGTTCTTACTCT 660  
QY 3715 CATGAAAGGCTGAAATGAAGGCACTATTCTCAACAAACCAAACTCATTTGTTCTG 3774  
Db 661 CACGAGAAAGTGAAATGAGGCACTCTCTCAGCANA-CANACTCATTTGTTCTG 719  
QY 3775 CAGGCAAAATGACCAACTGCTAAAA 3802  
Db 720 CAAG-CANAATGACCACTGCTAAAA 746

RESULT 11

LOCUS BG912161 843 bp mRNA linear EST 05-JUN-2001  
DEFINITION 602812833F1 NCI\_CGAP\_Brn67 Homo sapiens cdna clone IMAGE:4944657  
5', mRNA sequence.

ACCESSION BG912161  
VERSION BG912161.1 GI:14292637  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 843)  
NIH-MGC <http://mgi.nci.nih.gov/>

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Tissue Procurement: David N. Louis, M.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLAM10890 row: a column: 10  
High quality sequence stop: 778.

FEATURES  
source Location/Qualifiers

1..843  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4944657"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q  
loss"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Brn67"  
/note="Organ: Brain; Vector: pCMV-SPOrt6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

ORIGIN

Query Match 8.9%; Score 584.6; DB 12; Length 843;  
Best Local Similarity 92.8%; Pred. No. 6.9e-129;  
Matches 669; Conservative 0; Mismatches 44; Indels 8; Gaps 5;  
QY 4654 GAAAAAGCAGAGTGTGTAATCTCTTGGAAAACTCCCTGCTGAACTGGAAAGTGTAT 4713  
Db 122 GAAAAAGCAGAGTGTGTAATCTCTTGGAAAACTCCCTGCTGAACTGGAAAGTGTAT 181  
QY 4714 GACCGTCTAGACATGAATGACAGCTGCCCTTCAGTGCAGAGTGGTGTGGTGGCACC 4773  
Db 182 GACCGTCTAGACATGAATGACAGCTGCCCTTCAGTGCAGAGTGGTGTGGTGGCACC 241  
QY 4774 GAGAAAGGGCTCTACGCCCTGAAATGCTTTGAAAAAATCTCCCTAACCCATGTCCTCCAGGAAT 4833  
Db 242 GAGAAAGGGCTCTACGCCCTGAAATGCTTTGAAAAAATCTCCCTAACCCATGTCCTCCAGGAAT 301  
QY 4834 GGACGCTCTCCAAATTTATATCAAGACCTGAGAGCTGAGAGTACTCATGATAGCAGGA 4893  
Db 302 GGACGCTCTCCAAATTTATATCAAGACCTGAGAGCTGAGAGTACTCATGATAGCAGGA 361  
QY 4894 GAAGAGGGGCTCTGCTGCTTTGGAGCTGAAAGAACTGAAACAGTCCCTGGCCCTCAGTCC 4953  
Db 362 GAAGAGGGGCTCTGCTGCTTTGGAGCTGAAAGAACTGAAACAGTCCCTGGCCCTCAGTCC 421  
QY 4954 CACCTGCTGCCAGCCGCGACATCTCACCACAACTTTTGAAGCTGTCAAGGGCTGCCAC 5013  
Db 422 CACCTGCTGCCAGCCGCGACATCTCACCACAACTTTTGAAGCTGTCAAGGGCTGCCAC 481  
QY 5014 TTGTTTGGGCGAGCAAGATTGAAACGGGCTCTGCACTCTGTGCAGCCATGCCCGACAAA 5073  
Db 482 TTGTTTGGGCGAGCAAGATTGAAACGGGCTCTGCACTCTGTGCAGCCATGCCCGACAAA 541  
QY 5074 GTCCTCATTTCCGCTCAACGAAAACTCAGCAAAATCTGCATCTCCGAAAGAGATAGAG 5133  
Db 542 GTCCTCATTTCCGCTCAACGAAAACTCAGCAAAATCTGCATCTCCGAAAGAGATAGAG 601  
QY 5134 ACCTCAGAGCCCTGACGCTGTATCCACTTCACCAATTCACAGTATCTCTCAAT--GGAAACAA 5192  
Db 602 ACCTCAGAGCCCTGACGCTGTATCCACTTCACCAATTCACAGTATCTCTCAATTCATTGGGAACAA 661  
QY 5193 TAAATTTACG--AAATCGACATGAA--GCAGTACACGCTCGAGGAAT--CCTGGATAAGA 5248  
Db 662 TAAATTTACGAAATCGACATGAAGCAGTACACGCTCGAGGAATTTCTTGGATAACGA 721  
QY 5249 ATGACCATTTCTTGGCAGCTGCTGTG---TTTGGCGCTCTTCCACAGTCTCCCTGCTCT 5305  
Db 722 ATGACCATTTCTTGGCAGCTGCTGTG---TTTGGCGCTCTTCCACAGTCTCCCTGCTCT 781  
QY 5306 CAATCGTGCAGGTGAACAGCGCGAGGAGAGTACTTGTCTGTGTTTCCACGAAT 5365  
Db 782 AATCGTGCAGGTGAACAGCGCGAGGAGAGTATTGCTGTGTTTCCCTCCGACT 841  
QY 5366 T 5366  
Db 842 T 842

RESULT 12

LOCUS CF744580 647 bp mRNA linear EST 10-OCT-2003  
DEFINITION UI-M-GVO-clt-n-23-0-UI.r1 NIH\_BMAP\_GVO Mus musculus cdna clone  
IMAGE:30617710 5', mRNA sequence.  
ACCESSION CF744580  
VERSION CF744580.1 GI:37640920  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 647)  
NIH-MGC <http://mgi.nci.nih.gov/>  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: gqpbs-@mail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mousefl.html  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

FEATURES  
source

Seq primer: pYX-5,  
Location/Qualifiers  
1. 647  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30617710"  
/tissue\_type="whole brain"  
/dev\_stage="1,5, and 15 days newborn"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_GVO"  
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated with  
EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CGAAGTGAAT. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

ORIGIN

Query Match 8.7%; Score 573.6; DB 14; Length 647;  
Best Local Similarity 93.8%; Pred. No. 2.5e-126;  
Matches 608; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY	1149	CACCTCAAGTCGACGATGACACCTCCCAATTTTGATGAACAGAGAAGAAATTCGTGGGT	1208
DB	1	CACCTCAAGTCGACGATGACACCTCCCAATTTTGATGAACAGAGAAGAAATTCGTGGGT	60
QY	1209	TTCAATCCTCTCCGTGCGCAGCTGAGCCCTCTCAGGCTTCTCGGGTGAAGAACTCCGCTTGT	1268
DB	61	TTCAATCCTCTCTGTGCGCAGCTGAGCCCTCTCGGCTTCTCAGGCGAAGAGCTCGGCTTGT	120
QY	1269	GGGGTTTCTAGCAGACGACCTGGGATTTCTGTAGATCTGAGTCTGTGTCTCGG	1328
DB	121	GGGATTTCTGACAGACGACCTGGGATTTCTGTAGATCTGAGTCTGTGTCTCGG	180
QY	1329	TCGGAAGTCTCCCTCCCAAGACTAGCTCCATGGGAAAGAAATTTCTCATCAAAAGCAAGA	1388
DB	181	TCGGAAGTCTCCCTCCCAAGGTTAGCTCCATGGGAAAGAAATTTCTCATCAAAAGCAAGA	240
QY	1389	GCTACAGACTCTCAGGACACAGTGTACAGATGGAGCGAGGAATGACCGGTTACATCG	1448
DB	241	GTCTCAAGACTCTCCAGGACCAAGTGTACAGATGGAGCGAGGAATGACCGGTTACATCG	300
QY	1449	GAGAGTGTCTCAGAGTGGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCGCTCTGA	1508
DB	301	CAGAGTGTCTCAGAGTGGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCGCTCTGA	360
QY	1509	GACTCAGAGTCCCTCTCTGGAGCAGACCTTGCTACTACATCAGCAAGATGAGTAGCTT	1568
DB	361	GACTCAGAGTCCCTCTCTGGAGCAGACCTTGCTACTACATCAGCAAGATGAGTAGCTT	420
QY	1569	AAAGCGAAGTTTGGACGACCGATGGAGGTGTCACGAGGAGTACCAAGGACTGCA	1628

DB	421	AAAGCGAAGTTTGGACGACCGGATGGAGGTGTCACGAGGATGCAAAAGCTCTGCA	480
QY	1629	GCTTCTCATGATATCAGAGCAGACCGGAAAGCTCCAAAGAAATCAAAGAGAGGAGTA	1688
DB	481	GCTTCTCATGATATCAGAGCAGACCGGAAAGCTCCAAAGAAATCAAAGAGAGGAGTA	540
QY	1689	CCAGGCTCAAGTGGAAAGAAATGAGGTTGATGATGAATCAGTTGGAAGAGGATCTTGCTC	1748
DB	541	CCAGGCTCAAGTGGAGGAGATGAGGCTGATGAATCAGTTGGAAGAGGATCTTGCTC	599
QY	1749	AGCAAGAGACGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCG	1796
DB	600	AGCCCGCAGACGAGGATCTCTACGAATCTGAGCTGAGGAGTCTCG	647

RESULT 13  
CF723360  
LOCUS  
DEFINITION  
UI-M-GVO-cjh-j-20-0-UI.r1 NIH\_BMAP\_GVO Mus musculus cDNA clone  
IMAGE:30546187 5', mRNA sequence.  
CF723360  
ACCESSION  
VERSION  
KEYWORDS  
CF723360.1 GI:37597528  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC http://mgc.nci.nih.gov/  
1 (bases 1 to 730)  
AUTHORS  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: gqpbs-remail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mousefl.html  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

Seq primer: pYX-5,  
Location/Qualifiers  
1. 730  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30546187"  
/tissue\_type="whole brain"  
/dev\_stage="1,5, and 15 days newborn"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_GVO"  
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated with  
EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CGAAGTGAAT. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

ORIGIN

Query Match 8.7%; Score 572.8; DB 14; Length 730;  
Best Local Similarity 88.7%; Pred. No. 4.3e-126;

Matches 643; Conservative 0; Mismatches 75; Indels 7; Gaps 2;	
QY	5212 ATGAACGATACAGCTCGAGGAATTCCTGGATAGAAATGACATTCCTTGGACCTGCT 5271
Db	1 ATGAACGATACAGCTTCACGAGTTCTTGGACAAGACACATTCCTTGGACCTGCT 60
QY	5272 GTGTTTGGCGCTCTTCCAAACAGCTTCCCTGTCTCAATCGTCAGTGAACAGCGCAGG 5331
Db	61 GTGTTGGCTCTCTGTCCAAAGCTTCCCTGTCTCAATGTGACGGACAGCGCCGG 120
QY	5332 CAGCGAGGAGTACTGTGTGTTTCCACGAATTGGAGTGTTCGTGGAATCTTACCGA 5391
Db	121 CAGCGAGAGAAATACCTGTGTCTTCCACGAATTGGGGTGTTCGTGATCTTACCGA 180
QY	5392 AGACGTAGCGCACAGACATCTCAAGTGGAGTCCCTTACCTTGGCCCTTGGCTACAGA 5451
Db	181 AGACGTAGCGCACAGATGATCTCAAGTGGAGTCCCTTACCTTGGCCCTTGGCTACAGA 240
QY	5452 GAACCTATCTGTGTTGTGACCCACTTCAACTCACTCGAAGTAATGAGATCCAGCAGCG 5511
Db	241 GAACCTATCTGTGTTGTGACTCACTTCAACTCCCTGGGAATGATGAGATCCAGCAGCG 300
QY	5512 TCCTAGCAGGAGACCCCTCCGAGGTACCTGGACATCCGAAACCGCGGTACCTGGCG 5571
Db	301 TCCTAGCAGGAGACCCCTCCGAGCATATCTGGAATTCGAACCCCTCGCTACCTGGCG 360
QY	5572 CTGCGCAATTCCTCAGGAGCGATTTACTTGGCGTCTCATACAGGATAAATTAAGGTC 5631
Db	361 CCGCGAATTCCTCGGAGCGATTTACTTGGCGTCTCATACAGGATAAATTAAGGTC 420
QY	5632 ATTGCTGCAAGGAAACCTCTGTAAGAGTCCGCACTGAACACCCCGGGGCCGTC 5691
Db	421 ATATGCTGCAAGGAAACCTCTGTAAGAGTCCGCACTGAACACCCCGGGGCCGTC 480
QY	5692 ACCTCCGAGCAGCCCAACAGAGGAGCCCAACCCAGTACAGAGCAGATCACCAAG 5751
Db	481 ACCTCCGAGCAGCCCAACAGAGGAGCCCAACCAATACAGAGCAGATCACCAAG 540
QY	5752 CGCGTGGCTCCAGCCCGCGCCGCGAAGCCCGAGCCAGCCCGGAGAGCAGCACA 5811
Db	541 CGCGTGGCTCCAGCCCGCGCCGCGAAGCCCGAGCCAGCCCGGAGAGCAGCACA 600
QY	5812 CCCACCGCTAC-----GCGAGGGGGAGCCGAGCTGCGGAGGACAAGTCTCTGGC 5865
Db	601 CGCACCGCTACCGGAGCAGAGAGGGTCCGACAGAGCTGCGGAGGAGCAAGTCTCCAGC 660
QY	5866 CGCCCGCTGGAGCAGAGAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCC 5925
Db	661 CGCCCTGGAGCGGGAGAGT-CCGAGCGCGATGCTCAGCAGCTAGGAGAGCGGTCC 719
QY	5926 CCGGG 5930
Db	720 CAGGG 724
RESULT 14	
CD172337/c	
LOCUS	CD172337 826 bp mRNA linear EST 19-MAY-2003
DEFINITION	AGENCOURT_14063676 NIH_MGC_181 Homo sapiens cDNA clone
ACCESSION	IMAGE:30376012 5', mRNA sequence.
VERSION	CD172337
KEYWORDS	CD172337.1 GI:30853788
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	1. (bases 1 to 826)
COMMENT	NIH-MGC <a href="http://mgi.nci.nih.gov/">http://mgi.nci.nih.gov/</a>
	National Institutes of Health, Mammalian Gene Collection (MGC)
	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>

Tissue Procurement: Dr. Michael Brownstein	
cDNA Library Preparation: Invitrogen Corp	
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)	
found through the I.M.A.G.E. Consortium/LLNL at:	
<a href="http://image.llnl.gov">http://image.llnl.gov</a>	
Plate: NDAM437 row: P column: 05	
High quality sequence stop: 760.	
FEATURES	
source	Location/Qualifiers
1..826	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:30376012"
	/tissue_type="White Matter"
	/dev_stage="Unknown"
	/lab_host="DH10B-Ton A ( T1 and T5 phage resistances) "
	/clone_lib="NIH_MGC_181"
	/note="Vector: PCMV-SPO76.1; Site 1: NotI; Site 2: EcoRV
	(destroyed); Library is oligo-dT primed and directionally
	cloned (EcoRV site is destroyed upon cloning). Average
	insert size 1.42 Kb. Library was constructed by
	(Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN	
Query Match 8.7%; Score 570; DB 14; Length 826;	
Best Local Similarity 100.0%; Pred. No. 2.2e-125;	
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	6005 AGGGAAGAGGCGAGAGTGCCTCTCAAGTTTTCACGGTTAACTGCTATTTATGACT 6064
Db	740 AGGGAAGAGGCGAGAGTGCCTCTCAAGTTTTCACGGTTAACTGCTATTTATGACT 681
QY	6065 GGATATAAAGCTGGAGCAACCTGCCAGCTAACTGGTTCAGTCTCCTGAGGATCATCCAGTGA 6124
Db	680 GGATATAAAGCTGGAGCAACCTGCCAGCTAACTGGTTCAGTCTCCTGAGGATCATCCAGTGA 621
QY	6125 ATGAGAGAAATCCCGCAGCAGGTTGAAAGTCTGTTCTGAGAACAGATTAATGCTGACAG 6184
Db	620 ATGAGAGAAATCCCGCAGCAGGTTGAAAGTCTGTTCTGAGAACAGATTAATGCTGACAG 561
QY	6185 AGTTCAATGACTTCTAGACGTGGTGAATTAATAAATGGCTTAAGCTCGAGCGCAGC 6244
Db	560 AGTTCAATGACTTCTAGACGTGGTGAATTAATAAATGGCTTAAGCTCGAGCGCAGC 501
QY	6245 CACCTCTGCTTACAAAAGAGTCTTAGTGCACATGCTGTAAGAACAACTTGTAAACCC 6304
Db	500 CACCTCTGCTTACAAAAGAGTCTTAGTGCACATGCTGTAAGAACAACTTGTAAACCC 441
QY	6305 TCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCCGAGA 6364
Db	440 TCATCTAGAAATCAGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCCGAGA 381
QY	6365 GACAACTGTTGTTGATTTTGAAGGACAGGACCAACACCTGATTTAGTTCCATAGC 6424
Db	380 GACAACTGTTGTTGATTTTGAAGGACAGGACCAACACCTGATTTAGTTCCATAGC 321
QY	6425 CAGGCTTCAACAGGACCAAGTGGCTTGGCCCTTAAAAACACACAGATGACTGGAATGATGT 6484
Db	320 CAGGCTTCAACAGGACCAAGTGGCTTGGCCCTTAAAAACACACAGATGACTGGAATGATGT 261
QY	6485 GTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGGAGTTAGCATTTTGTG 6544
Db	260 GTGGCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGGAGTTAGCATTTTGTG 201
QY	6545 CTTAAGAAAAATCGAGAATGTAGGTTTAGA 6574
Db	200 CTTAAGAAAAATCGAGAATGTAGGTTTAGA 171
RESULT 15	
BQ421272/c	

LOCUS B0421272 894 bp mRNA linear EST 23-MAY-2002  
DEFINITION AGENCOURT 7761930 NIH\_MGC\_70 Homo sapiens cDNA clone IMAGE:6018187  
5', mRNA sequence.  
ACCESSION B0421272  
VERSION B0421272.1 GI:21116587  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 894)  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: sgapbs-x@mail.nih.gov  
Tissue Procurement: AICC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM13218 row: k column: 20  
High quality sequence stop: 720.  
Location/Qualifiers  
1. .894  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6018187"  
/tissue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 70"  
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."  
ORIGIN  
Query Match 8.7%; Score 570; DB 13; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.4e-125;  
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6005 AGGAGAGAGGCGAGAGTCCCTCTCAAGTTTTCACGGTTAACACTGTCACCTATTATGACT 6064  
DB 570 AGGAGAGAGGCGAGAGTCCCTCTCAAGTTTTCACGGTTAACACTGTCACCTATTATGACT 511  
QY 6065 GGAATATAAAGCTGGACAACTGCCAGCTAACTGGTTCAGTCTGAGGATCATCCAGCTGA 6124  
DB 510 GGAATATAAAGCTGGACAACTGCCAGCTAACTGGTTCAGTCTGAGGATCATCCAGCTGA 451  
QY 6125 ATGGAGAAATCCGGCAGCAGGTTGAAAAGTCTCTTCTGAGAACAGATTATTGCTGAGCAG 6184  
DB 450 ATGGAGAAATCCGGCAGCAGGTTGAAAAGTCTCTTCTGAGAACAGATTATTGCTGAGCAG 391  
QY 6185 AGTTTCATGTGACTCTAGAGCTGGTGCATTTAAATAATGGCTTAAGGCTGCAGAGCCAGC 6244  
DB 390 AGTTTCATGTGACTCTAGAGCTGGTGCATTTAAATAATGGCTTAAGGCTGCAGAGCCAGC 331  
QY 6245 CACCTCTGCTTACAAAAGAGTACTTAGTGCACATGACTGTAAAGAACAAATTGTAACCC 6304  
DB 330 CACCTCTGCTTACAAAAGAGTACTTAGTGCACATGACTGTAAAGAACAAATTGTAACCC 271  
QY 6305 TCATCTAGAAATCAGAAAGCTCTTAATTTCTATAGAAATGACACCTCCCTGGAGCCGAGA 6364  
DB 270 TCATCTAGAAATCAGAAAGCTCTTAATTTCTATAGAAATGACACCTCCCTGGAGCCGAGA 211  
QY 6365 GACAATCTGTTGTGATTTTGAAGGAGCAGGCAAGACCAACACTGTATTAGTTCCATAGC 6424  
DB 210 GACAATCTGTTGTGATTTTGAAGGAGCAGGCAAGACCAACACTGTATTAGTTCCATAGC 151  
QY 6425 CAGGCCTCAACAGGGACAAGTGGCTGGCCTTAAAAACACACAGATGACCTGGAAATGATGT 6484

Db 150 CAGGCCTCAACAGGGACAAGTGGCTGGCCTTAAAAACACACAGATGACTGGAATGATGT 91  
QY 6485 GTGGCCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGAGTAGTACATTCATTTTGG 6544  
Db 90 GTGGCCTCAGTCCCTGTTTCCAGAAATTTTACTGGCAAGAGTAGTACATTCATTTTGG 31  
QY 6545 CTTAAGAAAAATCGAGAATGTAGGTTTAGA 6574  
Db 30 CTTAAGAAAAATCGAGAATGTAGGTTTAGA 1  
Search completed: July 4, 2004, 07:39:38  
Job time : 14343 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2004, 10:07:00 ; Search time 104 Seconds  
(without alignment)  
5577.596 Million cell updates/sec

Title: US-10-017-216-2  
Perfect score: 10490  
Sequence: 1 MLKFKYGARNPLDAGAAEPI.....QLNGEIRQVEKSVLRDTC 2053

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Length	DB	ID	Description
1	10490	100.0	2053	5	AEE24079	Human MDP
2	10032.5	95.6	2054	5	AEE24150	Human kin
3	10032.5	95.6	2054	6	AAC26959	Human CRI
4	10022.5	95.5	2054	5	ABB81927	Human kin
5	10012	95.4	2053	4	AU03501	Human pro
6	10007	95.4	2066	5	ABG78362	Human pro
7	10007	95.4	2066	6	ADA05654	Human NOV
8	10005	95.4	2053	5	ABG78363	Human NOV
9	10005	95.4	2053	6	ADA05642	Human NOV
10	9656	92.0	2055	6	ABP97683	Polypepti
11	9656	92.0	2055	6	AAC26960	Human CRI
12	9487.5	90.4	1958	5	ABB81928	Human kin
13	7521	71.7	1619	7	ADD48584	Rat Prote
14	7521	71.7	1619	7	ADD46616	Rat Prote
15	6223.5	59.3	1286	3	AAB43359	Human ORF
16	6223.5	59.3	1286	6	AAC26961	Human CRI
17	6223.5	59.3	1286	7	ADD48586	Human Pro
18	6223.5	59.3	1286	7	ADD46618	Human Pro
19	4710.5	44.9	999	4	ABB11117	Human RHO
20	4519.5	43.1	940	7	ADD89967	Human can
21	4174.5	39.8	810	4	AU31443	Novel hum
22	4101	39.1	847	6	ADA05648	Human NOV
23	4005.5	38.2	832	6	ADA05646	Human NOV
24	3976.5	37.9	853	7	ADG09957	Novel pro
25	3155	30.1	623	6	ADA05644	Human NOV

26	2910	27.7	623	6	ADA05650	Human NOV
27	2892.5	27.6	638	6	ADA05652	Human NOV
28	2433	23.2	495	6	ABP97681	Amino aci
29	2433	23.2	497	6	ABP97687	Amino aci
30	2430	23.2	497	6	ABU10126	Novel hum
31	2430	23.2	497	7	AEE39504	Human kin
32	2425	23.1	497	5	AEE16261	Human kin
33	2417	23.0	497	6	ABU10127	Variant n
34	2155.5	20.6	494	6	ABP97682	Polypepti
35	1620	15.4	349	4	ABG15566	Novel hum
36	1446	13.8	1770	5	AEE25099	Human kin
37	1421.5	13.6	1719	5	AEE21707	Human PKI
38	1400	13.3	1738	7	ADE47738	Human NOV
39	1399.5	13.3	1664	7	ADE47740	Human NOV
40	1398.5	13.3	1797	4	ABG13880	Novel hum
41	1390	13.3	1728	3	AAB42069	Human ORF
42	1390	13.3	1728	6	AAB38363	Human ser
43	1382.5	13.2	1711	5	AAU97065	CDC42-bin
44	1382.5	13.2	1711	6	AAB38364	Human ser
45	1329	12.7	1252	4	ABB66357	Drosophil

## ALIGNMENTS

RESULT 1  
AAE24079  
XX AAE24079 standard; protein; 2053 AA.  
AC AAE24079;  
XX  
XX  
DT 04-OCT-2002 (first entry)  
XX  
DE Human MDPK protein.  
XX  
KW Human; myotonic dystrophy type protein kinase; MDPK; 13245 protein;  
KW tumorigenesis; tumour growth; tumour metastasis; viral infection;  
KW skeletal muscle disorder; muscular dystrophy; myotonic dystrophy;  
KW immune disorder; neoplastic disorder; gene therapy.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 53..303  
FT Domain /note= "pkinase domain"  
FT Domain 97..360  
FT Domain /note= "pkinase domain"  
FT Region 195..210  
FT Region /note= "Antigenic epitope"  
FT Region 217..229  
FT Region /note= "Serine/Threonine protein kinase active site signature"  
FT Region 455..475  
FT Domain /note= "Antigenic epitope"  
FT Domain 1568..1865  
FT Domain /note= "CNH domain"  
WO200234896-A2.  
02-MAY-2002.  
23-OCT-2001; 2001WO-US050636.  
23-OCT-2000; 2000US-0242429P.  
(MILL-) MILLENNIUM PHARM INC.  
Kapeller-Libermann R;  
WPI; 2002-479720/51.  
N-PSDB; AAD39191.  
Human myotonic dystrophy type protein kinase polypeptide and

PT polynucleotide useful for prognosticating, diagnosing, preventing or  
 PT inhibiting tumorigenesis, tumor growth, tumor metastasis and viral  
 XX infection.

PS Claim 8; Fig 3; 148pp; English.

XX The invention relates to human myotonic dystrophy type protein kinase  
 CC (MDPK) polypeptides designated as 13245 and nucleic acid molecules  
 CC encoding such polypeptides. 13245 molecules are used to develop  
 CC diagnostic and therapeutic agents for prognosticating, diagnosing,  
 CC preventing, inhibiting, alleviating or curing MDPK-related disorders.  
 CC Polypeptides of the invention are used to develop diagnostic and  
 CC therapeutic agents for 13245-mediated or related disorders such as  
 CC tumorigenesis, tumor growth, tumor metastasis, viral infection of a  
 CC cell, skeletal muscle disorders (e.g. muscular and myotonic dystrophies),  
 CC immune disorders and neoplastic disorders. The invention is also used in  
 CC gene therapy. The present sequence is human MDPK protein  
 XX

SQ Sequence 2053 AA;

Query Match	100.0%;	Score 10490;	DB 5;	Length 2053;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2053;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MLRFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALFVLFE	60		
DB 1	MLRFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPFMTQQQMSPLSREGILDALFVLFE	60		
QY 61	ECSQPALMKIKHVSFVRKYSUTIAELQELQPSAKDFEVRSLVGCCHFAEVQVREKATG	120		
DB 61	ECSQPALMKIKHVSFVRKYSUTIAELQELQPSAKDFEVRSLVGCCHFAEVQVREKATG	120		
QY 121	DIYAMKVMKKALLAQOVSFFBEERNILSRSTSPWIPOLQYAFQDKNHLYLMEEYQPGG	180		
DB 121	DIYAMKVMKKALLAQOVSFFBEERNILSRSTSPWIPOLQYAFQDKNHLYLMEEYQPGG	180		
QY 181	DLISLNRVEDQDENLIQFYLAELILAVHSVHLMGYVRDITKPNILVDRTGHIKLVDVF	240		
DB 181	DLISLNRVEDQDENLIQFYLAELILAVHSVHLMGYVRDITKPNILVDRTGHIKLVDVF	240		
QY 241	GSAAKNSNKNWNAKLPITGPDYMAPEVLTVMNGDGKTYGLDCDWSVGVIAIYEMIGR	300		
DB 241	GSAAKNSNKNWNAKLPITGPDYMAPEVLTVMNGDGKTYGLDCDWSVGVIAIYEMIGR	300		
QY 301	SPFAEGTSARTFNNINNFORFLKPPDDPKVSSDFDLIQSLLCGQKRELKFEGLCCHPFF	360		
DB 301	SPFAEGTSARTFNNINNFORFLKPPDDPKVSSDFDLIQSLLCGQKRELKFEGLCCHPFF	360		
QY 361	SKIDMWNIRNSPPFPVPTLKSDDTSNDFEPEKNSWVSSPQOLSPSGSGBELPVGFS	420		
DB 361	SKIDMWNIRNSPPFPVPTLKSDDTSNDFEPEKNSWVSSPQOLSPSGSGBELPVGFS	420		
QY 421	YSKALGTLGRSESVSGLDSPAKTSMEKKLITKSKELQSDQDKCHKMEQEMTRLHRRVS	480		
DB 421	YSKALGTLGRSESVSGLDSPAKTSMEKKLITKSKELQSDQDKCHKMEQEMTRLHRRVS	480		
QY 481	EVBAVLSQKEVELKASETQSLLEQDLATYITECSSLKRSLEQARMEVSQEDDKALQLLH	540		
DB 481	EVBAVLSQKEVELKASETQSLLEQDLATYITECSSLKRSLEQARMEVSQEDDKALQLLH	540		
QY 541	DIREQRKQEIKEQYQAOVEEMRLMMNQLEEDLVSAARRSDLYSELSRESRLAAEFK	600		
DB 541	DIREQRKQEIKEQYQAOVEEMRLMMNQLEEDLVSAARRSDLYSELSRESRLAAEFK	600		
QY 601	RKATECHQLLKAKQCKPVGVIYAKLEKINAEQQLKIQELQKLEKAAKEAERELEKL	660		
DB 601	RKATECHQLLKAKQCKPVGVIYAKLEKINAEQQLKIQELQKLEKAAKEAERELEKL	660		
QY 661	QNEDESSEGRKKLVAAEERHSLNKVRLTEWBERENRLKDDIOTKQQQIQQADKIL	720		
DB 661	QNEDESSEGRKKLVAAEERHSLNKVRLTEWBERENRLKDDIOTKQQQIQQADKIL	720		
QY 721	ELEKHEAQVSAQHLVHLKQEQHYEKKIVLDNQIKKDLADKETLENMMQREHEEAH	780		

DB 721	ELEKHEAQVSAQHLVHLKQEQHYEKKIVLDNQIKKDLADKETLENMMQREHEEAH	780
QY 781	EKGKILSEQAMINAMDSKIRSLRQIVELSEANKLAANSSLTQRMKAQZEMISBLRQ	840
DB 781	EKGKILSEQAMINAMDSKIRSLRQIVELSEANKLAANSSLTQRMKAQZEMISBLRQ	840
QY 841	QKFVLETQAGKLEAQNPKLEEQLEKISHQSDSKNRLLETLRLETVLSHEEQKLELKR	900
DB 841	QKFVLETQAGKLEAQNPKLEEQLEKISHQSDSKNRLLETLRLETVLSHEEQKLELKR	900
QY 901	QLTELOLSLOERESQTLQAARAALESQLRQAKTELETTAAEETTAQTAHRDEIQR	960
DB 901	QLTELOLSLOERESQTLQAARAALESQLRQAKTELETTAAEETTAQTAHRDEIQR	960
QY 961	KFDALRNSCTVITDLEEQNLQTNEDNAELNNQNFYLSKQLEDEASGANDEIVQLRSVDHL	1020
DB 961	KFDALRNSCTVITDLEEQNLQTNEDNAELNNQNFYLSKQLEDEASGANDEIVQLRSVDHL	1020
QY 1021	RREITEREMOLTSQKOTMEALKTTCTMLEEQVMDLEALNDELLEKQEWASRVLGDEK	1080
DB 1021	RREITEREMOLTSQKOTMEALKTTCTMLEEQVMDLEALNDELLEKQEWASRVLGDEK	1080
QY 1081	SQFCRVRELOQMLDTEKQSRADQIRITESQVVELAVKEHKAETILAQOALKQKLLKA	1140
DB 1081	SQFCRVRELOQMLDTEKQSRADQIRITESQVVELAVKEHKAETILAQOALKQKLLKA	1140
QY 1141	ESLSDKLNDLEKXHAMLENNARSLOQKLETERELKORLLEEQAKLQOQMDLQKNHIFRLT	1200
DB 1141	ESLSDKLNDLEKXHAMLENNARSLOQKLETERELKORLLEEQAKLQOQMDLQKNHIFRLT	1200
QY 1201	QGLQELADRALDKTERSSDLEYOLENIQVLYSHKVMMEGTISQOQTKLIDFLQAKMDQPA	1260
DB 1201	QGLQELADRALDKTERSSDLEYOLENIQVLYSHKVMMEGTISQOQTKLIDFLQAKMDQPA	1260
QY 1261	KKKKGLFSRKEDEPALPTQVPLQVNLKALKEKARCALEBALOKTRIELSAREEAA	1320
DB 1261	KKKKGLFSRKEDEPALPTQVPLQVNLKALKEKARCALEBALOKTRIELSAREEAA	1320
QY 1321	HRKATDHPHSTPATARQQIAMSALVSRPEHQPSAMSLAPPSSRRKESSTPEFSRRLK	1380
DB 1321	HRKATDHPHSTPATARQQIAMSALVSRPEHQPSAMSLAPPSSRRKESSTPEFSRRLK	1380
QY 1381	ERMHNIPHRFNVLNMRATKCAVCLDTVHFQASKLCQVWCHPKCSTCLPATCGLP	1440
DB 1381	ERMHNIPHRFNVLNMRATKCAVCLDTVHFQASKLCQVWCHPKCSTCLPATCGLP	1440
QY 1441	AEYATHTEAFCDKKNSSPGLQTKPESSSLHLEGWKMVPRNNKRGQGGWDRKYIVLEGSK	1500
DB 1441	AEYATHTEAFCDKKNSSPGLQTKPESSSLHLEGWKMVPRNNKRGQGGWDRKYIVLEGSK	1500
QY 1501	VLIYDNEAREAGQPVVEFELCIPDGDVSIHGAVGASELANATAKAEKADAKLGNL	1560
DB 1501	VLIYDNEAREAGQPVVEFELCIPDGDVSIHGAVGASELANATAKAEKADAKLGNL	1560
QY 1561	KLEGGDLDMNCTLPFSDQVVLVGTTEGLYALNVLNLSLTHVPGIGAVFQIYIKOLEKL	1620
DB 1561	KLEGGDLDMNCTLPFSDQVVLVGTTEGLYALNVLNLSLTHVPGIGAVFQIYIKOLEKL	1620
QY 1621	LMTAGEBALCLVDVKKVQSLQASHLPAQPDISPNIPEAVKCHLPFGAGKIENGLCICA	1680
DB 1621	LMTAGEBALCLVDVKKVQSLQASHLPAQPDISPNIPEAVKCHLPFGAGKIENGLCICA	1680
QY 1681	AMPSKVILIRYNELSKYCIKBTETSEPCSCIFHTNYSILIGTNKPYEDMKQYTLBEF	1740
DB 1681	AMPSKVILIRYNELSKYCIKBTETSEPCSCIFHTNYSILIGTNKPYEDMKQYTLBEF	1740
QY 1741	LDKNDHSLAPAVFAASSNSPFSIVQVNSAQREVEYLCPHFEGFVDSYGRSRTDCLK	1800
DB 1741	LDKNDHSLAPAVFAASSNSPFSIVQVNSAQREVEYLCPHFEGFVDSYGRSRTDCLK	1800
QY 1801	WSRLPLAFAYREPVLFTYHNSLEVIHQARSAGTAPARVLDIPNPRYLGPASSAIY	1860

```

Db 1801 WERLPLAFAYREPYLFTVTHFNSLEVIEIOARSSAGTPARAYLDIPNRYLQPAISSGAIY 1860
QY 1861 LASSYQDKLRVICCKGNLVKESGTEHHRGPTSRSSPNKRGPTTYNEHTKEVASSPAPP 1920
Db 1861 LASSYQDKLRVICCKGNLVKESGTEHHRGPTSRSSPNKRGPTTYNEHTKEVASSPAPP 1920
QY 1921 EGPSPHREPSTPHRYREGTELRDDKSGPRLERKSPGRLMSTRERSPGRLPFEDSSRG 1980
Db 1921 EGPSPHREPSTPHRYREGTELRDDKSGPRLERKSPGRLMSTRERSPGRLPFEDSSRG 1980
QY 1981 RLPAGAVRTPLSQVNVKGRGQSASQVFTNTVTYYDWNKKLDNLPAWNSVLRILQLNGEIR 2040
Db 1981 RLPAGAVRTPLSQVNVKGRGQSASQVFTNTVTYYDWNKKLDNLPAWNSVLRILQLNGEIR 2040
QY 2041 QQVEKSVLRDTC 2053
Db 2041 QQVEKSVLRDTC 2053

RESULT 2
AAE24150
ID AAE24150 standard; protein; 2054 AA.
AC AAE24150;
DT 23-SEP-2002 (first entry)
XX Human kinase (PKIN)-21 protein.
DE Human kinase (PKIN)-21 protein.
XX Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
KW development; hepatitis; cardiovascular; hypertension; drug screening;
KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
KW hypercholesterolemia; obesity; gene therapy; cytostatic; anti-HIV;
KW neuroprotective; hepatotropic; hypotensive; cardiant; nephrotropic;
XX Homo sapiens.
OS Homo sapiens.
XX
FH Key
FT Location/Qualifiers
FT 97..360
FT /note= "Eukaryotic protein kinase domain"
FT 98..241
FT /note= "Protein kinase domain"
FT 99..349
FT /note= "Protein kinase domain"
FT 101..241
FT /note= "Protein kinase domain"
FT 102..241
FT /note= "Protein kinase domain"
FT 249..349
FT /note= "Protein kinase domain"
FT 258..445
FT /note= "Protein kinase domain"
FT 258..349
FT /note= "Protein kinase domain"
FT 534..542
FT /note= "Domain found in NIK1-like kinase"
FT 544..875
FT /note= "Leucine zipper pattern"
FT 891..933
FT /note= "Domain found in NIK1-like kinase"
FT 964..975
FT /note= "Domain found in NIK1-like kinase"
FT 991..1012
FT /note= "Leucine zipper pattern"
FT 1015..1067
FT /note= "Domain found in NIK1-like kinase"
FT 1057..1078
FT /note= "Leucine zipper pattern"
FT 1159..1180

```

```

FT /note= "Leucine zipper pattern"
FT 1217..1255
FT /note= "Domain found in NIK1-like kinase"
FT 1388..1434
FT /note= "Domain found in NIK1-like kinase"
FT Binding-site
FT 1390..1438
FT /note= "Phorbol esters/diacylglycerol binding site"
FT Binding-site
FT 1403..1466
FT /note= "Phorbol esters/diacylglycerol binding site"
FT Domain
FT 1471..1590
FT /note= "PH domain"
FT Domain
FT 1619..1916
FT /note= "CNH domain"
FT Domain
FT 1759..1802
FT /note= "Domain found in NIK1-like kinase"
FT Domain
FT 1819..1831
FT /note= "Domain found in NIK1-like kinase"
FT Domain
FT 1851..1880
FT /note= "Domain found in NIK1-like kinase"
XX
XX WO200233099-A2.
XX
XX 25-APR-2002.
XX
XX 20-OCT-2001; 2001WO-US047728.
XX
XX 20-OCT-2000; 2000US-0242410P.
XX 27-OCT-2000; 2000US-0244068P.
XX 03-NOV-2000; 2000US-0245708P.
XX 09-NOV-2000; 2000US-0247672P.
XX 16-NOV-2000; 2000US-0249565P.
XX 22-NOV-2000; 2000US-0252730P.
XX 01-DEC-2000; 2000US-0250807P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Gururajan R, Baughn MR, Wallia NK, Elliott VS, Xu Y, Arvizu C;
XX Yac MG, Rankumar J, Ding L, Tang YT, Hafalia AJA, Nguyen DB;
XX Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;
XX Recipon SA, Lu DAW, Borowsky ML, Thornton M, Swarnaker A;
XX Thangavelu K, Khan FA, Ison CH;
XX
XX WPI; 2002-454603/48.
XX N-PSDB; AAD38864.
XX
XX New human kinase polypeptide, for diagnosing, preventing and treating
XX cancer, immune system disorders, growth and development disorders,
XX cardiovascular disorders and lipid disorders.
XX
XX Claim 1; Page 177-182; 210pp; English.
XX
XX The invention relates human kinases (PKIN) and their corresponding
XX nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
XX treating and preventing cancer, an immune system disorder (e.g., acquired
XX immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,
XX atherosclerosis, multiple sclerosis, psoriasis), disorders affecting
XX growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),
XX cardiovascular disorder (e.g., hypertension, myocardial infarction,
XX Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,
XX Gaucher's disease, Niemann-Pick's disease, hypercholesterolemia,
XX hyperlipidaemia, obesity), and for assessing the effects of exogenous
XX compounds. Anti-PKIN antibody is useful in a diagnostic test for a
XX condition or a disease associated with the expression of PKIN in a
XX biological sample. A composition comprising PKIN or an agonist or
XX antagonist of PKIN is useful for treating a disease or condition
XX associated with decreased or increased expression of functional PKIN.
XX PKIN is useful in a number of drug screening techniques and to analyse
XX the proteome of a tissue or cell type. PKIN DNA is useful for creating
XX knockin humanised animals or transgenic animals to model human diseases,
XX and in somatic or germline gene therapy. The present sequence is human
XX PKIN protein
XX
XX Sequence 2054 AA;
SQ

```

Query Match	95.6%;	Score 10032.5;	DB 5;	Length 2054;	
Best Local Similarity	95.8%;	Pred. No. 0;			
Matches 1981;	Conservative	2;	Mismatches	4;	Indels 81; Gaps 3;
Qy	1	MLRFKYGARPLDAGAAEPIASRASRLNLPFOGKPPFMTQQQWSPISREGILDALFVLFE	60		
Db	1	MLRFKYGARPLDAGAAEPIASRASRLNLPFOGKPPFMTQQQWSPISREGILDALFVLFE	60		
Qy	61	ECQOPALMKIKHVSNFVKYSDTIIAELOELQPSAKDFEVRSLVGCCHFAEVOVVRKATG	120		
Db	61	ECQOPALMKIKHVSNFVKYSDTIIAELOELQPSAKDFEVRSLVGCCHFAEVOVVRKATG	120		
Qy	121	DIYAMKVMKKALLAQEQVSFFBEERNILSRSTSPWIPOLQYAFQDKNHLHLMEEYOPGG	180		
Db	121	DIYAMKVMKKALLAQEQVSFFBEERNILSRSTSPWIPOLQYAFQDKNHLHLMEEYOPGG	180		
Qy	181	DLISLNLRYEDQDENLIQSVLAELILAVHSVHLMGVVHRDIKPENILVDRTGHIKLVD	240		
Db	181	DLISLNLRYEDQDENLIQSVLAELILAVHSVHLMGVVHRDIKPENILVDRTGHIKLVD	240		
Qy	241	GSAAKMNSNKNVNAKPIGTPDYMAPEVLTVMNGDGKGYGLBCDWMVSGVVIAYEMIYGR	300		
Db	241	GSAAKMNSNKNVNAKPIGTPDYMAPEVLTVMNGDGKGYGLBCDWMVSGVVIAYEMIYGR	300		
Qy	301	SPAEGTSARTFNINMNFQFLFPPDDPKVSSDFLDLIQSLLCGQKRLKFEGLCCHPFF	360		
Db	301	SPAEGTSARTFNINMNFQFLFPPDDPKVSSDFLDLIQSLLCGQKRLKFEGLCCHPFF	360		
Qy	361	SKIDWNINRSPPPFVPTLKSDDDTNFDPEKNVWVSSPCQLSPGSGSGEELPVGFS	420		
Db	361	SKIDWNINRSPPPFVPTLKSDDDTNFDPEKNVWVSSPCQLSPGSGSGEELPVGFS	420		
Qy	421	YSKALGTLGRSESVVSGLDSPAKTSSMEKLLIKSKLEQDSQDKCHQWQEMTRLHRRVS	480		
Db	421	YSKALGTLGRSESVVSGLDSPAKTSSMEKLLIKSKLEQDSQDKCHQWQEMTRLHRRVS	480		
Qy	481	EVHAVISQKEVELKASPTORSLLQDLATYITCSSLKRSLEQARMEVSGEDDKALQLLH	540		
Db	481	EVHAVISQKEVELKASPTORSLLQDLATYITCSSLKRSLEQARMEVSGEDDKALQLLH	540		
Qy	541	DIREQSRKLEIKQEYQCAQVEMRLMMNQLBEDLVSAARRSDLYSESLRESLAAEFPK	600		
Db	541	DIREQSRKLEIKQEYQCAQVEMRLMMNQLBEDLVSAARRSDLYSESLRESLAAEFPK	600		
Qy	601	RKATECOHLLKAKDQGPVGYAKLEKINASQQLKIQLQSKLAKA-----	648		
Db	601	RKATECOHLLKAKDQGPVGYAKLEKINASQQLKIQLQSKLAKAVKASTEATELLQ	650		
Qy	649	-----AKERAERELEKLNREDSSEGIRKKLVEAERRHSLSENKVLETMERENRLKOD	704		
Db	661	NIQAKERAPARELEKLNREDSSEGIRKKLVEAERRHSLSENKVLETMERENRLKOD	720		
Qy	705	IQTQSQOIQWADKILELEKPHREAOVSAQHLEVHLKQEQHYEELIKVLJDNQIKKDLAD	764		
Db	721	IQTQSQOIQWADKILELEKPHREAOVSAQHLEVHLKQEQHYEELIKVLJDNQIKKDLAD	780		
Qy	765	KEPLENMQRHEEAHEKGIKLEQKAMINAMDSKTRLEQRIVELSEANKLAANSLSFT	824		
Db	781	KEPLENMQRHEEAHEKGIKLEQKAMINAMDSKTRLEQRIVELSEANKLAANSLSFT	840		
Qy	825	QRNKAAQEMISLRQOKFYLETQAGKLEAQNKLEBQLEKISHQHSDKVRILLELETRL	884		
Db	841	QRNKAAQEMISLRQOKFYLETQAGKLEAQNKLEBQLEKISHQHSDKVRILLELETRL	900		
Qy	885	REVSLEHEEQKLEKROLTELOZLSQRESQLTALQAARAALSQLRQAQTELEETTAEA	944		
Db	901	REVSLEHEEQKLEKROLTELOZLSQRESQLTALQAARAALSQLRQAQTELEETTAEA	960		
Qy	945	EEBIQALTARHDEIQRFDAIRNCTVITDLEBQNLQLTEDNAELNNQNYFLSKQDLEAS	1004		
Db	961	EEBIQALTARHDEIQRFDAIRNCTVITDLEBQNLQLTEDNAELNNQNYFLSKQDLEAS	1020		

RESULT 3  
AA026959  
ID AAO26959 standard; protein; 2054 AA.  
XX AC AAO26959;  
XX DT 01-MAY-2003 (first entry)  
XX DE Human CR1K protein sequence, SEQ ID No 2.  
XX KW Anorectic; hypotensive; cardiatic; antilipaemic; cerebroprotective;  
XX KW antigout; osteopathic; antiarthritic; cytosatic; antidepressant;  
XX KW immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;  
XX KW neuroprotective; antiinflammatory; antidiabetic; analgesic;  
XX KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;  
XX KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;  
XX KW central nervous system disorder; chronic obstructive pulmonary disease;  
XX KW diabetes; pain.  
XX OS Homo sapiens.  
XX PN WO2003004523-A1.  
XX PD 16-JAN-2003.  
XX PF 28-JUN-2002; 2002WO-BP007156.  
XX PR 02-JUL-2001; 2001US-0301841P.  
XX PR 11-DEC-2001; 2001US-0338651P.  
XX PR 25-APR-2002; 2002US-0375014P.  
XX PA (FARB ) BAYER AG.  
XX PI Zhu Z;  
XX WPI: 2003-221576/21.  
XX DR N-PSDB; AAL55214.  
XX PT New human citron rho/rac-interacting kinase (CR1K) polypeptide and  
XX PT polynucleotide, useful in preventing, ameliorating or treating diseases  
XX PT associated with human CR1K dysfunction, e.g. obesity, diabetes or  
XX PT Alzheimer's disease.  
XX PS Claim 1; Fig 2; 237pp; English.  
XX CC The invention relates to an isolated polynucleotide encoding a human  
XX CC citron rho/rac-interacting kinase polypeptide. The isolated  
XX CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the  
XX CC specification. The human citron rho/rac-interacting kinase (CR1K)  
XX CC polypeptide and polynucleotide are useful in preventing, ameliorating, or  
XX CC treating diseases associated with human CR1K dysfunction such as obesity  
XX CC and obesity-associated comorbidities (e.g. hypertension, coronary artery  
XX CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of  
XX CC cancer including endometrial, breast, prostate and colon cancer),  
XX CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood  
XX CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's  
XX CC disease), chronic obstructive pulmonary disease, or diabetes. These can  
XX CC also be used to treat pain associated with the disorders. The human CR1K  
XX CC polypeptide is also useful in diagnostic assays or in genetic testing.  
XX CC The expression vector or the reagent is useful in preparing a medicament  
XX CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a  
XX CC central nervous system disorder, or chronic obstructive pulmonary  
XX CC disease. The fusion protein is useful for generating antibodies against a  
XX CC CR1K polypeptide and for use in various assay systems. The methods are  
XX CC useful in producing and detecting the polynucleotide and polypeptide and  
XX CC in screening for agents that modulate the activity of the human CR1K  
XX CC invention.  
XX SQ Sequence 2054 AA;  
Query Match 95.6%; Score 10032.5; DB 6; Length 2054;  
Best Local Similarity 95.8%; Pred. No. 0;

Matches 1981; Conservative 2; Mismatches 4; Indels 81; Gaps 3;  
QY 1 MLKFKYGARNPLDAGAAEPTASRASRLNLFQGGPPFTQOQMSPLREGILDALFVLFE 60  
DB 1 MLKFKYGARNPLDAGAAEPTASRASRLNLFQGGPPFTQOQMSPLREGILDALFVLFE 60  
QY 61 ECSQPALMKIKHVSNFVRKYSDTTIAELQELQPSAKDFEVRSLVCGGHFAEVVVREKATG 120  
DB 61 ECSQPALMKIKHVSNFVRKYSDTTIAELQELQPSAKDFEVRSLVCGGHFAEVVVREKATG 120  
QY 121 DIYAMKWKKALLAQOQVFFEBERNILSRSTSPWIPQOYAFODKXNHLVMEYOPGG 180  
DB 121 DIYAMKWKKALLAQOQVFFEBERNILSRSTSPWIPQOYAFODKXNHLVMEYOPGG 180  
QY 181 DLLSLLNRYEDQDENLIQFYLAELIILAVHSVHLMGVVRDIDKPNILVDTGHIKLVDF 240  
DB 181 DLLSLLNRYEDQDENLIQFYLAELIILAVHSVHLMGVVRDIDKPNILVDTGHIKLVDF 240  
QY 241 GSAAKMNSNMVNAKLPIGTPDYNAPEVLTVMNGDGKGTGGLDCDWMSVGVIAEMVYGR 300  
DB 241 GSAAKMNSNMVNAKLPIGTPDYNAPEVLTVMNGDGKGTGGLDCDWMSVGVIAEMVYGR 300  
QY 301 SPFAEGTSARTFNINMNFQFLKFPDDPKYSSDFDLIQLSLCCQKERLKFEGLCCHPFF 360  
DB 301 SPFAEGTSARTFNINMNFQFLKFPDDPKYSSDFDLIQLSLCCQKERLKFEGLCCHPFF 360  
QY 361 SKIDWNIRNSPPFPVPTLKSDDTSNFDPEPKNSWVSSPCQLSPSGFSGEELPFVGF 420  
DB 361 SKIDWNIRNSPPFPVPTLKSDDTSNFDPEPKNSWVSSPCQLSPSGFSGEELPFVGF 420  
QY 421 YSKALGILGRSESVVGLDGPACTSSMEKLLIKSKELQSDQKCHKWEQMTLHRVS 480  
DB 421 YSKALGILGRSESVVGLDGPACTSSMEKLLIKSKELQSDQKCHKWEQMTLHRVS 480  
QY 481 EVEAVLSQKEVELKASQTSRLLEQDLATVITECSSLKRSLEQARMEVSQEDDKALQLH 540  
DB 481 EVEAVLSQKEVELKASQTSRLLEQDLATVITECSSLKRSLEQARMEVSQEDDKALQLH 540  
QY 541 DIRQSKRLQIKQEQYCAQVEENRMLMNOLEEDLVARRSDIYSELRESRLAAEFK 600  
DB 541 DIRQSKRLQIKQEQYCAQVEENRMLMNOLEEDLVARRSDIYSELRESRLAAEFK 600  
QY 601 RKATECOHKLLKAKDQKPEVGEYAKLEKINABOOLKIQLQOEKLEKA----- 648  
DB 601 RKATECOHKLLKAKDQKPEVGEYAKLEKINABOOLKIQLQOEKLEKAVKASTATELLQ 660  
QY 649 ----AKPRAERELEKLNREDSSSGIRKKLVEAEERHSHLENKYKRLTWERRENRLKDD 704  
DB 661 NIRAQKRAERELEKLNREDSSSGIRKKLVEAEERHSHLENKYKRLTWERRENRLKDD 720  
QY 705 IQTSQOIQQMADKILEEKHREAOVSACHLEVHLKQKQHYEEKIKVLDNQIKKDLAD 764  
DB 721 IQTSQOIQQMADKILEEKHREAOVSACHLEVHLKQKQHYEEKIKVLDNQIKKDLAD 780  
QY 765 KETLENMWQRHEEAEHKGKILSQKAMINAMDSKIRSLQRIVELSEANKLAANSFLT 824  
DB 781 KETLENMWQRHEEAEHKGKILSQKAMINAMDSKIRSLQRIVELSEANKLAANSFLT 840  
QY 825 QRNKAQEMISELRQOKFYLETQAGKLEAQNKRKLEQLEKISHQSDSKNRLLELTRL 884  
DB 841 QRNKAQEMISELRQOKFYLETQAGKLEAQNKRKLEQLEKISHQSDSKNRLLELTRL 900  
QY 885 REVLSLEHEEKLEKROLTELQLSQERESQTLQAARAALESQLRQAKTELETTAAE 944  
DB 901 REVLSLEHEEKLEKROLTELQLSQERESQTLQAARAALESQLRQAKTELETTAAE 960  
QY 945 EBEIQAALTAHRDEIQRFKDALRNSCTVITDLEEQNLQTDENAEELNNQNFYLSQOLDEAS 1004  
DB 961 EBEIQAALTAHRDEIQRFKDALRNSCTVITDLEEQNLQTDENAEELNNQNFYLSQOLDEAS 1020  
QY 1005 GANDEIVQLRSEVDHLRREITEREMQLTSSOKQTEALKTCTMLEEQVMDLEALNDELLE 1064  
DB 1021 GANDEIVQLRSEVDHLRREITEREMQLTSSOKQTEALKTCTMLEEQVMDLEALNDELLE 1080

```
QY 1065 KERQWEAWRSVLGDEKSPQFCRVREIQRMLDTEKQSRARADQBITESRQVVELAVKEHKA 1124
Db 1081 KERQWEAWRSVLGDEKSPQFCRVREIQRMLDTEKQSRARADQBITESRQVVELAVKEHKA 1140
QY 1125 EIALQALKEQKLKABSLDKLNDLEKHAMLEMNARSIOOKLETERELKORLLEBOAK 1184
Db 1141 EIALQALKEQKLKABSLDKLNDLEKHAMLEMNARSIOOKLETERELKORLLEBOAK 1200
QY 1185 LQQQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYQLENIQVLYSHEKVMQETISQ 1244
Db 1201 LQQQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYQLENIQVLYSHEKVMQETISQ 1260
QY 1245 QTKLIDFLQAKMPPAKKKGLFSRRKEDPALFTQVPLQYNELKLALEKEKARCABLEEA 1304
Db 1261 QTKLIDFLQAKMPPAKKKK-----VPLQYNELKLALEKEKARCABLEEA 1305
QY 1305 LQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSAIVRSPEHQPSAMSLAPPSS 1364
Db 1306 LQKTRIELRSARBEAAHRKATDHPHSTPATARQQIAMSAIVRSPEHQPSAMSLAPPSS 1365
QY 1365 RRKESSTPEFSRRLKERMHHNIPHRPNVGLNMRATKCAVCLDTVHFGRQAQKCLECQVM 1424
Db 1366 RRKESSTPEFSRRLKERMHHNIPHRPNVGLNMRATKCAVCLDTVHFGRQAQKCLECQVM 1425
QY 1425 CHPKCSCTCLPATCGLPAEYATHTFAFCRDKNNSPGLQTKPESSSLHLBGWKNVPRNNKR 1484
Db 1426 CHPKCSCTCLPATCGLPAEYATHTFAFCRDKNNSPGLQTKPESSSLHLBGWKNVPRNNKR 1485
QY 1485 GQQGWRKYIYVLEGSKVLYVDNFAEAGQRPVEPELCLPDGQVSHGAVGASELANTAK 1544
Db 1486 GQQGWRKYIYVLEGSKVLYVDNFAEAGQRPVEPELCLPDGQVSHGAVGASELANTAK 1545
QY 1545 A-----EKAADAKL 1554
Db 1546 ADVPYILKMESHPTTCWPORTULYLAFFDPKQWWTALSVVAGGRVSRKADAKL 1605
QY 1555 LGNSLLKLEGGDRDLNMCTLPFSDQVVLVGTBEGLYALNVKNSLTHVPGIGAVFIYII 1614
Db 1606 LGNSLLKLEGGDRDLNMCTLPFSDQVVLVGTBEGLYALNVKNSLTHVPGIGAVFIYII 1665
QY 1615 KDLEKLIMINAGERALCLVDVKKVQSLAQSHLPAQPDISPNIPEAVKCHLFGAGKIBN 1674
Db 1666 KDLEKLIMINAGERALCLVDVKKVQSLAQSHLPAQPDISPNIPEAVKCHLFGAGKIBN 1725
QY 1675 GLGICAMPKSVILRYNENLSKYCRKETETSEPCSCIHFTNYSILIGTNKFYEIDMKQ 1734
Db 1726 GLGICAMPKSVILRYNENLSKYCRKETETSEPCSCIHFTNYSILIGTNKFYEIDMKQ 1785
QY 1735 YTLLEFLDKNDHSLAPFAVFAASNSFPVSIVQNSAGQREYLLCFHEFGVFDVSYGRS 1794
Db 1786 YTLLEFLDKNDHSLAPFAVFAASNSFPVSIVQNSAGQREYLLCFHEFGVFDVSYGRS 1845
QY 1795 RTDGLKWSRLPLAFAYREPVLFTVTHNSLEVIETIQARSAGTPARAYLDIPNPRVLGPAT 1854
Db 1846 RTDGLKWSRLPLAFAYREPVLFTVTHNSLEVIETIQARSAGTPARAYLDIPNPRVLGPAT 1905
QY 1855 SSGAIYLAASYQDKLVI CCKGNLVKESGTEHHRGPSTSRSSPNKRGPPTYNEHITKRYA 1914
Db 1906 SSGAIYLAASYQDKLVI CCKGNLVKESGTEHHRGPSTSRSSPNKRGPPTYNEHITKRYA 1965
QY 1915 SSPAPEGSHPEPSTPHRYREGRTELARDKSPGRPLERKSPGMLSTREERSPGRLF 1974
Db 1966 SSPAPEGSHPEPSTPHRYREGRTELARDKSPGRPLERKSPGMLSTREERSPGRLF 2025
QY 1975 EDSSRGLPAGAVRTPLSQVKNRGQSA 2002
Db 2026 EDSSRGLPAGAVRTPLSQVKNWQDSS 2053
```

RESULT 4  
ABB81927  
ID ABB81927 standard; protein; 2054 AA.

```
XX ABB81927;
XX AC
XX DT 10-OCT-2002 (first entry)
XX DE Human kinase #1.
XX KW Human; kinase; enzyme; serine-threonine kinase; nootropic; cytostatic;
XX KW Citron rho-interacting kinase; Gene therapy; mental disorder; cancer.
XX OS Homo sapiens.
XX PN WO200259325-A2.
XX PD 01-AUG-2002.
XX PF 20-DEC-2001; 2001WO-US050497.
XX PR 27-DEC-2000; 2000US-0258335P.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Yu X, Miranda M, Fridde CJ;
XX DR WPI; 2002-599796/64.
XX DR N-PSDB; ABQ78870.
XX PT Novel polynucleotide encoding human proteins that are structurally
XX PT similar to animal kinases, useful for drug screening, diagnosis, in gene
XX PT therapy of disorders and diseases e.g. cancer and pharmacogenomic
XX PS applications.
XX PS Claim 2; Page 39-43; 50pp; English.
XX CC The invention relates to a novel human protein that shares structural
XX CC similarity with animal kinases, including serine-threonine kinases,
XX CC particularly Citron rho-interacting kinases. The proteins of the
XX CC invention have nootropic and cytostatic activity. The polynucleotides may
XX CC have a use in gene therapy. The encoded novel polypeptides are useful for
XX CC generating antibodies, as reagents in diagnostic assays, for identifying
XX CC other cellular gene products related to NHP and as reagents in assays for
XX CC screening for compounds that are useful in the treatment of mental,
XX CC biological or medical disorders and diseases including cancer. The
XX CC sequence represents a novel human kinase of the invention
XX SQ Sequence 2054 AA;
Query Match 95.5%; Score 10022.5; DB 5; Length 2054;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1979; Conservative 3; Mismatches 5; Indels 81; Gaps 3;
QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPMTQQQMSPLSREGILDALFVIFE 60
Db 1 MLKFKYGARNPLDAGAAEPIASRASRLNLFQGGKPPMTQQQMSPLSREGILDALFVIFE 60
QY 61 ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVSRLVGCCHFAEVSQVVRKATG 120
Db 61 ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVSRLVGCCHFAEVSQVVRKATG 120
QY 121 DIYAMKWKKKALIAQOVSVFFEEERNILSRSTSPWIPQLQYAFQDNKHLVLMEEYQPGG 180
Db 121 DIYAMKWKKKALIAQOVSVFFEEERNILSRSTSPWIPQLQYAFQDNKHLVLMEEYQPGG 180
QY 181 DLLSLNRYEDQDENLIQFYLAELILAVSHVLMGVHRCIKPENTILVDRTGHILKLVDF 240
Db 181 DLLSLNRYEDQDENLIQFYLAELILAVSHVLMGVHRCIKPENTILVDRTGHILKLVDF 240
QY 241 GSAAKMNSKNVNAKLPIGTDPYMAPEVLTVMNGDGKGTGVLDCDWMWSGVIAIYMYGR 300
Db 241 GSAAKMNSKNVNAKLPIGTDPYMAPEVLTVMNGDGKGTGVLDCDWMWSGVIAIYMYGR 300
QY 301 SPFAEGTSATFTNNIMNFQRLKFPDDPKVSSQFPLDIQSLLCQKRLKFEGLCCHPFF 360
Db 301 SPFAEGTSATFTNNIMNFQRLKFPDDPKVSSQFPLDIQSLLCQKRLKFEGLCCHPFF 360
```

Db 301 SFAEGTSARTFNNIMNFRLFKFPDDPKVSDFDLDLIQSLICGQKRLKFEGLCCHPFF 360  
Qy 361 SKTDWNNIRNSPPFPVPTT.KSDDDTNPNFDEPEKNISWSSPQCLSPSGSGSEBELPVGFS 420  
Db 361 SKTDWNNIRNSPPFPVPTT.KSDDDTNPNFDEPEKNISWSSPQCLSPSGSGSEBELPVGFS 420  
Qy 421 YSKALGILGRSESVSGSLDPAKTSMEKKLLIKSKELQSDQDKCHKMQEOTRHLHRRVS 480  
Db 421 YSKALGILGRSESVSGSLDPAKTSMEKKLLIKSKELQSDQDKCHKMQEOTRHLHRRVS 480  
Qy 481 EYAVLSQKEVELKASQTSRSLLEQDLATVITECSSLKRSLEQARVEVSQEDDKALQLLH 540  
Db 481 EYAVLSQKEVELKASQTSRSLLEQDLATVITECSSLKRSLEQARVEVSQEDDKALQLLH 540  
Qy 541 DIREQSRKQETKEQSYQAVQEMRLMMNQLDEDLVSARRSDLYESELRESRLAABEFK 600  
Db 541 DIREQSRKQETKEQSYQAVQEMRLMMNQLDEDLVSARRSDLYESELRESRLAABEFK 600  
Qy 601 RKATECOHLLKAKQDQKPEVGEYAKLEKINAEQQLKIQELQEKLEKA----- 648  
Db 601 RKATECOHLLKAKQDQKPEVGEYAKLEKINAEQQLKIQELQEKLEKAVKASTEATELLQ 660  
Qy 649 ----AKERARELEKLNQNRDESEGIKKLVABERRHSLNKKVLETMERENRLKDD 704  
Db 661 NINQAKERARELEKLNQNRDESEGIKKLVABERRHSLNKKVLETMERENRLKDD 720  
Qy 705 IQTKSQIQOMADKILBEKEKREAOVSQAHLEVLHKKQEQHYEBKIKVLDNOIKKDLAD 764  
Db 721 IQTKSQIQOMADKILBEKEKREAOVSQAHLEVLHKKQEQHYEBKIKVLDNQIKKDLAD 780  
Qy 765 KETLENNMQRHEEAEHEKKILSEQKAMINAMDSKIRSLRQIRVELSEANKLAANSUFT 824  
Db 781 KETLENNMQRHEEAEHEKKILSEQKAMINAMDSKIRSLRQIRVELSEANKLAANSUFT 840  
Qy 825 QRNMAQEMISELROCKFYLETQAGKLEAQRNKLBEQLEKISHODHSDKNRLLIELETRL 884  
Db 841 QRNMAQEMISELROCKFYLETQAGKLEAQRNKLBEQLEKISHODHSDKNRLLIELETRL 900  
Qy 885 REVSLHEHEQKLELKRQLTELQSLQERESQLTALQAAFAALESQJRAKTELETTTAA 944  
Db 901 REVSLHEHEQKLELKRQLTELQSLQERESQLTALQAAFAALESQJRAKTELETTTAA 960  
Qy 945 EBEIQALTAHDEIQKFPALNSCTVITDLBEQNLQTLDEADNNAELNNQNFYLSKQLEAS 1004  
Db 961 EBEIQALTAHDEIQKFPALNSCTVITDLBEQNLQTLDEADNNAELNNQNFYLSKQLEAS 1020  
Qy 1005 GANDEIVQLRSEVDHLRRIITEREMOLTSQKQTMELKTTCTMLEEQVMDLEALNDELLE 1064  
Db 1021 GANDEIVQLRSEVDHLRRIITEREMOLTSQKQTMELKTTCTMLEEQVMDLEALNDELLE 1080  
Qy 1065 KERQWEAHSVLGDEKQSECFVREHLOQMLDTEKQSRAPADQIRITESQOVVELAVKEHKA 1124  
Db 1081 KERQWEAHSVLGDEKQSECFVREHLOQMLDTEKQSRAPADQIRITESQOVVELAVKEHKA 1140  
Qy 1125 EILALQOALKQKLEKAEISLSDKLNLEKHAMLENNARSLOQKLETERELKORLLEEQAK 1184  
Db 1141 EILALQOALKQKLEKAEISLSDKLNLEKHAMLENNARSLOQKLETERELKORLLEEQAK 1200  
Qy 1185 LQQMDLQKNHIFRLTQGLQEALEDRADLLKTERSDLEYOLENIQVLYSHEKVMEGTTSQ 1244  
Db 1201 LQQMDLQKNHIFRLTQGLQEALEDRADLLKTERSDLEYOLENIQVLYSHEKVMEGTTSQ 1260  
Qy 1245 QTKLIDFLQAKMDQPAKKKGULFSRRKEDPALPTQVPLQYNELKLALKEKARCALBEA 1304  
Db 1261 QTKLIDFLQAKMDQPAKKKGULFSRRKEDPALPTQVPLQYNELKLALKEKARCALBEA 1305  
Qy 1305 LQKTRIELRSAREEAHRAKTDHPHSTPATARQIAMSALVRSPEHQPSAMSLILAPSS 1364  
Db 1306 LQKTRIELRSAREEAHRAKTDHPHSTPATARQIAMSALVRSPEHQPSAMSLILAPSS 1365  
Qy 1365 RRKESSTPEEFRRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTHFGRQASKCLECQVM 1424  
Db 1366 RRKESSTPEEFRRRLKERMHHNIPHRFNVGLNMRATKCAVCLDTHFGRQASKCLECQVM 1425

Qy 1425 CHPKCSTCLPATCGLPAEYATHTEATCRDKQNSFGLQTKPESSSLHLEGMKVPNNKR 1484  
Db 1426 CHPKCSTCLPATCGLPAEYATHTEATCRDKQNSFGLQTKPESSSLHLEGMKVPNNKR 1485  
Qy 1485 GOQGWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANAK 1544  
Db 1486 GOQGWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANAK 1545  
Qy 1545 A-----EKAEBADKL 1554  
Db 1546 ADVFYILKMESHPTTTCWPGRTLYLLAPSPDKQWVTALESVVAGGRVSREKAEADAKL 1605  
Qy 1555 LGNSLLKLEGGDRDLDMNCTLPFSQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYII 1614  
Db 1606 LGNSLLKLEGGDRDLDMNCTLPFSQVVLVGTTEGLYALNVLKNSLTHVPGIGAVFOIYII 1665  
Qy 1615 KDLKELMIAGEERALCLVDVKKYKSLAQSHLPAQDIPNIFEAVKGGCHLFGAGKIEN 1674  
Db 1666 KDLKELMIAGEERALCLVDVKKYKSLAQSHLPAQDIPNIFEAVKGGCHLFGAGKIEN 1725  
Qy 1675 GLCTCAAMPSKVILRYNENLSKYCIKEIETSEPCSCIHFTNYSIILGTNKFYEIDMKQ 1734  
Db 1726 GLCTCAAMPSKVILRYNENLSKYCIKEIETSEPCSCIHFTNYSIILGTNKFYEIDMKQ 1785  
Qy 1735 YTLBEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREBEYLLCFHEFGVFDVSYGRRS 1794  
Db 1786 YTLBEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREBEYLLCFHEFGVFDVSYGRRS 1845  
Qy 1795 RTDILKWSRLPLAPAYREPYLFTVTHFNSLEVIETIQARSSAGTTPARAVLDIPNRYLGPAL 1854  
Db 1846 RTDILKWSRLPLAPAYREPYLFTVTHFNSLEVIETIQARSSAGTTPARAVLDIPNRYLGPAL 1905  
Qy 1855 SSGAIYILASSYQDQLRVIQCKGNLVKESGTEHHRGPGSTSSSPNKRGPPTVNEHITKRVA 1914  
Db 1906 SSGAIYILASSYQDQLRVIQCKGNLVKESGTEHHRGPGSTSSSPNKRGPPTVNEHITKRVA 1965  
Qy 1915 SSPAPPEGPHPREPSTPHRYREGRETELRRDKSPGRELREKSPGRMLSTRERSPORLF 1974  
Db 1966 SSPAPPEGPHPREPSTPHRYREGRETELRRDKSPGRELREKSPGRMLSTRERSPORLF 2025  
Qy 1975 EDSRGRPLPAGAVRTPLSQVNGKRGOSA 2002  
Db 2026 EDSRGRPLPAGAVRTPLSQVNGKRGOSA 2053

RESULT 5  
AAU03501  
ID AAU03501 standard; protein; 2053 AA.  
XX  
AC AAU03501;  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Human protein kinase #1.  
XX  
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
KW metabolic disorder; immune related disease; neurological disorder;  
KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
KW reproductive disorder.  
OS Homo sapiens.  
XX  
FN WO200138503-A2.  
XX  
PD 31-MAY-2001.  
XX  
PF 22-NOV-2000; 2000WO-US032085.  
XX  
PR 24-NOV-1999; 99US-0167482P.  
XX  
PA (SUGE-) SUGEN INC.  
XX

PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
PI Flanagan P, Clary D;  
XX WPI; 2001-343950/36.  
DR N-PSDB; AAS06701.

XX Nucleic acids encoding human kinase polypeptides, useful for preventing  
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
PT neuronal-associated diseases, and microbial infections.

XX Claim 7; Fig 2; 433pp; English.

PS CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel  
CC protein kinases have been identified as members of the tyrosine or  
CC serine/threonine kinase (PTK and STK) families. The polynucleotides  
CC encoding protein kinases and the polypeptides may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC cancers (especially cancers of haematopoietic origin), cardiovascular  
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
CC immune related diseases (e.g. rheumatoid arthritis), neurological  
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
CC Additionally, polynucleotides encoding protein kinases may be used for  
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase  
CC polypeptides may be used as antigens in the production of antibodies  
CC against the protein kinases and in assays to identify modulators of  
CC protein kinase expression and activity

XX SQ Sequence 2053 AA;

Query Match 95.4%; Score 10012; DB 4; Length 2053;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 1979; Conservative 2; Mismatches 5; Indels 82; Gaps 4;  
QY 1 MLKFKYGARNPLDAGAAEPTASSASRLNLFQFKPPFMTQQQSPLSRSGILDALFVLFE 60  
DB 1 MLKFKYGARNPLDAGAAEPTASSASRLNLFQFKPPFMTQQQSPLSRSGILDALFVLFE 60  
QY 61 ECQOPALMKIKHVSNFVRK-YSOTIAELQELQSAKDFFVRSVLGCGHFAEVQVREKAT 119  
DB 61 ECQOPALMKIKHVSNFVPEVYSTIAELQELQSAKDFFVRSVLGCGHFAEVQVREKAT 120  
QY 120 GDYAMKVMKKALLAQEQVSPFEERNILSRSTSPWIPOLQVAFODKNHLYIMBYQPG 179  
DB 121 GDYAMKVMKKALLAQEQVSPFEERNILSRSTSPWIPOLQVAFODKNHLYIMBYQPG 180  
QY 180 GDILSLINRYEDQDENLQFYLAELTILAVSHVLMGYVHRDIPKENILVDRTHIKLVD 239  
DB 181 GDLSLLNRYEDQDENLQFYLAELTILAVSHVLMGYVHRDIPKENILVDRTHIKLVD 240  
QY 240 FGSAAKVNSKVMNAKLPICTPDYMAPEVLTVNMGDKGTGYGLDCDWSVGVYAYEMIYG 299  
DB 241 FGSAAKVNSKVMNAKLPICTPDYMAPEVLTVNMGDKGTGYGLDCDWSVGVYAYEMIYG 300  
QY 300 RSPFAECTSARTNNINMFORLKFDDPKVSSDFDLTQSLCCGKERLKEGLCCHPF 359  
DB 301 RSPFAECTSARTNNINMFORLKFDDPKVSSDFDLTQSLCCGKERLKEGLCCHPF 360  
QY 360 FSKIDMNNIRNSPPFPVPTLKSDDDTSNFDPEPKNSWSSSPCQLSPGSGSEELPFVGF 419  
DB 361 FSKIDMNNIRNSPPFPVPTLKSDDDTSNFDPEPKNSWSSSPCQLSPGSGSEELPFVGF 420  
QY 420 SYSKALCILGRSSVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHMEQWTRLHRRV 479  
DB 421 SYSKALCILGRSSVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHMEQWTRLHRRV 480  
QY 480 SEVEAVLSQKEVELKASQETORSILLEODLATYITECSSLRSLPQARMEVSQEDDKALQLL 539  
DB 481 SEVEAVLSQKEVELKASQETORSILLEODLATYITECSSLRSLPQARMEVSQEDDKALQLL 540  
QY 540 HDTREOSRKLQETKEQBYQAVQVEMRLMMNQLEEDLVASRRSDLYSELSRSLAAEEF 599

DB HDIREQSRKLQETKEQBYQAVQVEMRLMMNQLEEDLVASRRSDLYSELSRSLAAEEF 600  
QY KRKATECHKLLKAKDQKPEVGEYAKLEKINAEQOLKIQELQEKLEKA----- 648  
DB KRKATECHKLLKAKDQKPEVGEYAKLEKINAEQOLKIQELQEKLEKAVKATEATELL 660  
QY -----AKERAERELEKLQNRDSDSEGIRKKLVBAEERRHSHLENKVKLETWERRNRLKD 703  
DB QNIRQAKERAERELEKLQNRDSDSEGIRKKLVBAEERRHSHLENKVKLETWERRNRLKD 720  
QY DIQTKSQOIQOMADKILLEKEKHREAVSAHLVHLKQKEQHYVEEKIKVDNQIKDILA 763  
DB DIQTKSQOIQOMADKILLEKEKHREAVSAHLVHLKQKEQHYVEEKIKVDNQIKDILA 780  
QY DKETLENMQRHHEEBAHEKGKILSEQAMINAMDSKIRSLQRTIVELSEANKLAANSLF 823  
DB DKETLENMQRHHEEBAHEKGKILSEQAMINAMDSKIRSLQRTIVELSEANKLAANSLF 840  
QY TORNMKAQEEEMISELRQOKFYLETQAGKLEAQNRLKEEQLEKISHQDHSQKRLLEETR 883  
DB TORNMKAQEEEMISELRQOKFYLETQAGKLEAQNRLKEEQLEKISHQDHSQKRLLEETR 900  
QY LREVSLEHEBEQKLEKQLTELQLSLOERSQLTALQAAARAALESQLRQAKTELEETAE 943  
DB LREVSLEHEBEQKLEKQLTELQLSLOERSQLTALQAAARAALESQLRQAKTELEETAE 960  
QY AEEBIIQALTARDEIQRFKDALRNSCTVITDLBEEQLNQLTEDNAELNNQNFYLSKQLEA 1003  
DB AEEBIIQALTARDEIQRFKDALRNSCTVITDLBEEQLNQLTEDNAELNNQNFYLSKQLEA 1020  
QY SGANDEIVQLRSEVDHLRREITEREMOLTSQKQTMELKTTCTMLBEQVMDLALNDELL 1063  
DB SGANDEIVQLRSEVDHLRREITEREMOLTSQKQTMELKTTCTMLBEQVMDLALNDELL 1080  
QY EKERQWEANWSVLGDEKSOFEQVRELRQMLDTEKQSRARADQRIITBSRQVVELAVKEHK 1123  
DB EKERQWEANWSVLGDEKSOFEQVRELRQMLDTEKQSRARADQRIITBSRQVVELAVKEHK 1140  
QY AEIILAQALKEQKLAESLSDKLNLEKXHAMLENNARSLQOKLETERELKORLLEEQ 1183  
DB AEIILAQALKEQKLAESLSDKLNLEKXHAMLENNARSLQOKLETERELKORLLEEQ 1200  
QY KLOQOMDLQKNHIFRLTQGLQELADRADLLKTBRSDEYOLENIQVLYSHEKVMQEGTIS 1243  
DB KLOQOMDLQKNHIFRLTQGLQELADRADLLKTBRSDEYOLENIQVLYSHEKVMQEGTIS 1260  
QY QOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLALEKEKARCABLEE 1303  
DB QOTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLALEKEKARCABLEE 1305  
QY ALQKTRIELRSARBEAAHRAKATDHPHSTPATATQOIAVSAIVRSPHQPSAMSLAPPS 1363  
DB ALQKTRIELRSARBEAAHRAKATDHPHSTPATATQOIAVSAIVRSPHQPSAMSLAPPS 1365  
QY SRRKESSTPEFBSRRLKERMHNNIPHRFNVGLNMRATKCAVCLDTVHFGQASKCLECQV 1423  
DB SRRKESSTPEFBSRRLKERMHNNIPHRFNVGLNMRATKCAVCLDTVHFGQASKCLECQV 1425  
QY MCHPKCSTCLPATCGLPAEYATHTFEAFCDKXNSPGLOTKEPSSSLHLGSMKVPNNK 1483  
DB MCHPKCSTCLPATCGLPAEYATHTFEAFCDKXNSPGLOTKEPSSSLHLGSMKVPNNK 1485  
QY RGQOGWRDKYIVLEGSKVLYVDNEAREAGORPVEEFELCLPDGDSVTHGAVGASELANTA 1543  
DB RGQOGWRDKYIVLEGSKVLYVDNEAREAGORPVEEFELCLPDGDSVTHGAVGASELANTA 1545  
QY KA-----EKAEADAK 1553  
DB KADVPYILKMESHPHHTCWPGRITLYLLAPFPDQKRWVTALESVYAGRVSRKAEADAK 1605  
QY LLGSHLLKLEGGDRLDNCTLPFSDQVVLVGTBEGLVALNVLKNSLTHVPQIGAVFOIYI 1613

Db 1606 LLGNSLLKLEGGDRDLDMNCTLPFDQVLVGTGTEGLYALNVLKNSLTHVFGIGAVFIYI 1665  
QY 1614 IKDLEKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDISNIPEAVKGCHLFGAGKIE 1673  
Db 1666 IKDLEKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDISNIPEAVKGCHLFGAGKIE 1725  
QY 1674 NGLICIAAMPKSVILRYNENISKYCIKIEIETSPSCSIHFTNYSILIGTNKFIYIDMK 1733  
Db 1726 NGLICIAAMPKSVILRYNENISKYCIKIEIETSPSCSIHFTNYSILIGTNKFIYIDMK 1785  
QY 1734 QYTLLEFLDKHSLAPAVFAASSNFPVSIYQVNSAGOREEYLLCFHFGYFVDSYGR 1793  
Db 1786 QYTLLEFLDKHSLAPAVFAASSNFPVSIYQVNSAGOREEYLLCFHFGYFVDSYGR 1845  
QY 1794 SRTDDLKWSRLPLAFAYREPFLVTHFNLSLEVIEIQARSSAGTAPARAVLDIPNRYLQPA 1853  
Db 1846 SRTDDLKWSRLPLAFAYREPFLVTHFNLSLEVIEIQARSSAGTAPARAVLDIPNRYLQPA 1905  
QY 1854 ISSGAIYLASSYQDKLRVLCCKGNLVKESGTEHHRGPTSRSPNKRGPPTVNEHITKRV 1913  
Db 1906 ISSGAIYLASSYQDKLRVLCCKGNLVKESGTEHHRGPTSRSPNKRGPPTVNEHITKRV 1965  
QY 1914 ASSPAPPEGSHPRPSTPHRYREGTELRDKSPGRPLERKSPGRMLSTRERSPGRL 1973  
Db 1966 ASSPAPPEGSHPRPSTPHRYREGTELRDKSPGRPLERKSPGRMLSTRERSPGRL 2025  
QY 1974 FDSRGRPLPAGAVRTPLSQVNGRQGS 2001  
Db 2026 FDSRGRPLPAGAVRTPLSQVNGRQHS 2053

RESULT 6  
ABG78362  
ID ABG78362 standard; protein; 2066 AA.  
XX  
AC ABG78362;  
XX  
DT 15-NOV-2002 (first entry)  
DE Human protein, homologous to kinases, designated NOV3a.  
XX  
KW Human; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;  
KW atherosclerosis; diabetes; cell signalling; metabolic pathway;  
KW cellular receptor; downstream effector; cancer; gene therapy;  
KW hypertension; congenital heart defect; aortic stenosis; obesity;  
KW infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;  
KW neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;  
KW haematopoietic disease; scleroderma; fertility; immunogen;  
KW idiopathic thrombocytopenic purpura; graft versus host disease;  
KW Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;  
KW systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;  
KW stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;  
KW pain; alcoholism; transgenic.  
OS Homo sapiens.  
XX  
PN WO200226826-A2.  
XX  
PD 04-APR-2002.  
XX  
PF 27-SEP-2001; 2001WO-US042336.  
XX  
PR 27-SEP-2000; 2000US-0235631P.  
PR 27-SEP-2000; 2000US-0235633P.  
PR 27-SEP-2000; 2000US-0235808P.  
PR 27-SEP-2000; 2000US-0236064P.  
PR 27-SEP-2000; 2000US-0236065P.  
PR 27-SEP-2000; 2000US-0236066P.  
PR 28-SEP-2000; 2000US-0236135P.  
PR 03-OCT-2000; 2000US-0237434P.  
PR 06-OCT-2000; 2000US-0238321P.  
PR 06-OCT-2000; 2000US-0238396P.  
PR 06-OCT-2000; 2000US-0238399P.

16-MAR-2001; 2001US-0276667P.  
31-MAY-2001; 2001US-0294823P.  
12-JUL-2001; 2001US-0304868P.  
26-SEP-2001; 2001US-00235831.  
XX (CURA-) CURAGEN CORP.  
PA Gerlach VL, Macdougall JR, Smithson G, Millet I, Stone D;  
PI Gunther E, Ellerman K, Grose WM, Alsobrook JP, Lepley DM;  
PI Burgess CE, Padigar M, Kekuda R, Spytek KA, Leach MD, Shimkets RA;  
XX WPI; 2002-499860/53.  
DR N-PSDB; ABS63435.  
XX Novel isolated NOVX polypeptides and polynucleotides homologous to  
PT attractin, plexin-like family of proteins, useful for treating  
PT atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and  
PT stroke.  
XX Claim 1; Page 42; 308pp; English.  
CC The invention discloses the isolated human polypeptides, and  
CC polynucleotides encoding them, that have been designated NOVX. The  
CC polypeptides, polynucleotides and antibodies are useful in treating or  
CC preventing a NOVX-associated disorder which is cardiomyopathy,  
CC atherosclerosis and diabetes in a human, where the disorder is related to  
CC cell signal processing and metabolic pathway modulation. They can also be  
CC used in determining the presence of, or predisposition to, a disease  
CC associated with altered levels of the polypeptides and polynucleotides of  
CC any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for  
CC identifying an agent that binds to, or that modulates the expression or  
CC activity of the polypeptide, for identifying an agent which is cellular  
CC receptor or downstream effector, for treating or preventing a NOVX-  
CC associated disorder and as a pharmaceutical composition comprising the  
CC polypeptide, polynucleotide or the antibody. The polypeptides and  
CC polynucleotides are useful in diagnostic applications (e.g. as a marker  
CC for cancerous cells or tissue types) where their amounts are assessed, or  
CC for the manufacture of a medicament (e.g. gene therapy) for treating or  
CC preventing disorders or syndromes such as hypertension, congenital heart  
CC defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,  
CC Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,  
CC haemophilia, dyslipidemia, haematopoietic diseases, scleroderma,  
CC fertility, idiopathic thrombocytopenic purpura, graft versus host  
CC diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune  
CC disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,  
CC allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar  
CC ataxia, pain and alcoholism. They may also be used as immunogens to  
CC produce antibodies specific for the invention, and as vaccines.  
CC Transgenic cells containing a NOVX expressing construct are useful to  
CC produce non-human transgenic animals for studying the function and/or  
CC activity of the NOVX proteins and for identifying and/or evaluating  
CC modulators of NOVX protein activity. Transgenic cells containing a NOVX  
CC expressing construct are useful to produce non-human transgenic animals  
CC for studying the function and/or activity of the NOVX proteins and for  
CC identifying and/or evaluating modulators of NOVX protein activity. The  
CC sequences presented in ABG78359-ABG78371 are the human NOV1-NOV8 proteins  
XX Sequence 2066 AA;

Query Match 95.4%; Score 10007; DB 5; Length 2066;  
Best Local Similarity 95.6%; Pred. No. 0;  
Matches 1979; Conservative 3; Mismatches 7; Indels 82; Gaps 4;  
QY 1 MLKFYKGYARNPLDAGAAEPTASRASRLNLFQCKPFPMTQQQMSPLSRGILDALFVLFE 60  
Db 1 MLKFYKGYARNPLDAGAAEPTASRASRLNLFQCKPFPMTQQQMSPLSRGILDALFVLFE 60  
QY 61 ECSQPALMKIKHVSNFVRKYSYDTIAELQELQPSAKDFEVRSLVGCGHFAEVVREKATG 120  
Db 61 ECSQPALMKIKHVSNFVRKYSYDTIAELQELQPSAKDFEVRSLVGCGHFAEVVREKATG 120  
QY 121 DIYAMKMKKALLAQEQNSPFEEENILSRSPHPIQLOYAFODKXHLHLYMEEVQPGG 180

121 DIYAMVKKKALLAQEQVSPFEERNILSRSTSPWIPQLQYAFQDKNHLXLYLWMEYQPGG 180  
181 DLLSLNRYEDQDENLQFYLAELILAVHSVHLMGYVHRDIPKPNILVDRGTGHKLVDF 240  
181 DLLSLNRYEDQDENLQFYLAELILAVHSVHLMGYVHRDIPKPNILVDRGTGHKLVDF 240  
241 GSAKMNKNVNAKLPIGTDPDYMAREVLTVMNGDGKTYGLDCDMSVGVIAIYGR 300  
241 GSAKMNKNK-VNAKLPIGTDPDYMAREVLTVMNGDGKTYGLDCDMSVGVIAIYGR 299  
301 SPAEGTSARTFNINMNFORFLKFPDDPKVSSDFDLIOGLCGQKERLKFEGLCCHPFF 360  
300 SPAEGTSARTFNINMNFORFLKFPDDPKVSSDFDLIOGLCGQKERLKFEGLCCHPFF 359  
361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDPEPEKNISWSSPCQLSPSGFSGBELPVGFS 420  
360 SKIDWNIRNAPPPFPVPTLKSDDDTSNFDPEPEKNISWSSPCQLSPSGFSGBELPVGFS 419  
421 YSALGILGSESWSGLSPAKTSWEKKLIIKSKELQSDQKCHKMEQENTRLHRRVS 480  
420 YSALGILGSESWSGLSPAKTSWEKKLIIKSKELQSDQKCHKMEQENTRLHRRVS 479  
481 EVEAVLSQKEVELKASSETORSLEQDLATYITECSSLSKRSLEQARMEVSQEDDKALQLLH 540  
480 EVEAVLSQKEVELKASSETORSLEQDLATYITECSSLSKRSLEQARMEVSQEDDKALQLLH 539  
541 DIREQSRKLOEIKOEQYQAVEMRLMMQLEBDIVSARRSDLYSELSERSLAABEPK 600  
540 DIREQSRKLOEIKOEQYQAVEMRLMMQLEBDIVSARRSDLYSELSERSLAABEPK 599  
601 RKATECOHLLKAKDQKPVGVYAKLEKINABQQLKIOELOKLEKA----- 648  
600 RKATECOHLLKAKDQKPVGVYAKLEKINABQQLKIOELOKLEKAVKASTATELQLQ 659  
649 ----AKERAERELEKQNRDSESEGIKKLVEAERHRSLENNKVKLETMERENRLKDD 704  
660 NIRAQERAERELEKQNRDSESEGIKKLVEAERHRSLENNKVKLETMERENRLKDD 719  
705 IQTKSQOIQMAQKILILEBKHEBAQVSAQHLEVHLKQEQHYEKKIVLDNQIKKDLAD 764  
720 IQTKSQOIQMAQKILILEBKHEBAQVSAQHLEVHLKQEQHYEKKIVLDNQIKKDLAD 779  
765 KETLENMQRHEBEAHEKGIKLSQKAMINAMDSKIRSLQRIVELSEANKLAANSSLT 824  
780 KETLENMQRHEBEAHEKGIKLSQKAMINAMDSKIRSLQRIVELSEANKLAANSSLT 839  
825 QRMKAQEMISLROQKFYLETQAGKLAQNRKLEQLEKISHQDHSQNRLLLETEL 884  
840 QRMKAQEMISLROQKFYLETQAGKLAQNRKLEQLEKISHQDHSQNRLLLETEL 899  
885 REVSLSHEEQKLEKQLTELQLSQERESQALQAARAALESQLRQAQTELETTABA 944  
900 REVSLSHEEQKLEKQLTELQLSQERESQALQAARAALESQLRQAQTELETTABA 959  
945 EEHIIQALTARHDIQKFDALRNSCTVITDLEBQLNQLTEDNAELNNQNFYLSKQLEAS 1004  
960 EEHIIQALTARHDIQKFDALRNSCTVITDLEBQLNQLTEDNAELNNQNFYLSKQLEAS 1019  
1005 GANDEIVQLRSEVDHLRREITEREMOLTSQKQMEALKTTCTMLEBQVMDLEALNDELLE 1064  
1020 GANDEIVQLRSEVDHLRREITEREMOLTSQKQMEALKTTCTMLEBQVMDLEALNDELLE 1079  
1065 KERQWEAWRSVLGDEKSQFCRVRLEQRLMDTEKQSRARADQRIITESROVVELAVKEHKA 1124  
1080 KERQWEAWRSVLGDEKSQFCRVRLEQRLMDTEKQSRARADQRIITESROVVELAVKEHKA 1139  
1125 EILALQALKEQKLAESLSDKLNLEKKHAMLENNARSLOQKLETERELKORLLEEQAK 1184  
1140 EILALQALKEQKLAESLSDKLNLEKKHAMLENNARSLOQKLETERELKORLLEEQAK 1199  
1185 LOQOMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1244  
1200 LOQOMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYOLENIQVLYSHEKVMGTTISQ 1259

1245 QTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALPTQVFPLOYNELKALAEKAKCAELEEA 1304  
1260 QTKLIDFLQAKMDQPAKKKK-----VPLQYNELKALAEKAKCAELEEA 1304  
1305 LQKTRIELRSAREBAARHKATDHPHPSTPATARQOIAWSAIVSRPEHQPSAMSLAPPSS 1364  
1305 LQKTRIELRSAREBAARHKATDHPHPSTPATARQOIAWSAIVSRPEHQPSAMSLAPPSS 1364  
1365 RRKESSTPEEFSSRLKERMHNIIPHRENVGLNMRATCAVCLDTVHFGRQASKCLEQVM 1424  
1365 RRKESSTPEEFSSRLKERMHNIIPHRENVGLNMRATCAVCLDTVHFGRQASKCLEQVM 1424  
1425 CHPKCSTCLPATCGLPABYATHFTEAFCDKMNSPGLQTEPSSSLHLEGMMKVPRNKR 1484  
1425 CHPKCSTCLPATCGLPABYATHFTEAFCDKMNSPGLQTEPSSSLHLEGMMKVPRNKR 1484  
1485 GQQGWRKYIVLEGSKVLIYDNEAREAGQRPVEFEFELCLPDGDVSIHGAVGASELANTAK 1544  
1485 GQQGWRKYIVLEGSKVLIYDNEAREAGQRPVEFEFELCLPDGDVSIHGAVGASELANTAK 1544  
1545 A-----EKAADAKL 1554  
1545 ADVPIILKMESHPTTCWPGRTLYLLAPSPDKQWVTALESVVAGGRVSRKAEADAKL 1604  
1555 LGNSLLKLEGDDRLDMNCTLPFSDQVVLVGTTEGLYALNVLXSLTHVPGIGAVFOIYI 1614  
1605 LGNSLLKLEGDDRLDMNCTLPFSDQVVLVGTTEGLYALNVLXSLTHVPGIGAVFOIYI 1664  
1615 KDLKILMIAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKVGCHLFGAGKIEN 1674  
1665 KDLKILMIAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIPEAVKVGCHLFGAGKIEN 1724  
1675 GLCICAMPKSVVILRYNENISKYCIKREIETSEPCSIHFTNYSLIGTNKFYEIDMKQ 1734  
1725 GLCICAMPKSVVILRYNENISKYCIKREIETSEPCSIHFTNYSLIGTNKFYEIDMKQ 1784  
1735 YTLSEFELKXNDHSLAPAVFAASNSFPVSIQVNSAQGQREYLLCPHEFGVFDVSYGRRS 1794  
1785 YTLSEFELKXNDHSLAPAVFAASNSFPVSIQVNSAQGQREYLLCPHEFGVFDVSYGRRS 1844  
1795 RTDDLKWSRLPLAPAYREPVLVTHFNSLEVIHQARSSAGTARAYLDIPNRYLGPAL 1854  
1845 RTDDLKWSRLPLAPAYREPVLVTHFNSLEVIHQARSSAGTARAYLDIPNRYLGPAL 1904  
1855 SSGAIIYLASSYQDKRLVICCKGNLVKESGTEHHRGTPSTSRSSPNKRGPTTYNEHITKXVA 1914  
1905 SSGAIIYLASSYQDKRLVICCKGNLVKESGTEHHRGTPSTSRSSPNKRGPTTYNEHITKXVA 1964  
1915 SSPAPPSPGPHRPPSTPHRYREGRTLRDKSPGRPLEREKSPGRMLSTRRSPGRPLF 1974  
1965 SSPAPPSPGPHRPPSTPHRYREGRTLRDKSPGRPLEREKSPGRMLSTRRSPGRPLF 2024  
1975 BSSRGRPLPAGAVETPLSQVNKRGQASOV 2005  
2025 BSSRGRPLPAGAVETPLSQVNKRGQASOV 2055

## RESULT 7

ADA05654  
ID ADA05654 standard; protein; 2066 AA.

XX ADA05654;

XX 06-NOV-2003 (first entry)

XX Human NOVlg protein SEQ ID NO:14.

XX human; NOVlg; antidiabetic; anorectic; antibacterial; virucide;

XX immunomodulator; cytosolic; nootropic; neuroprotective;

XX antiparkinsonian; antidiabetic; gene therapy; human disease;

XX metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

KW	immune disorder; haematopoietic disorder; dyslipidaemia.
XX	
OS	Homo sapiens.
XX	
PN	WO2003029424-A2.
XX	
PD	10-APR-2003.
XX	
PF	02-OCT-2002; 2002WO-US031373.
XX	
PR	02-OCT-2001; 2001US-0326483P.
PR	05-OCT-2001; 2001US-0327435P.
PR	05-OCT-2001; 2001US-0327449P.
PR	09-OCT-2001; 2001US-0327917P.
PR	09-OCT-2001; 2001US-0328029P.
PR	09-OCT-2001; 2001US-0328044P.
PR	09-OCT-2001; 2001US-0328056P.
PR	12-OCT-2001; 2001US-0328849P.
PR	15-OCT-2001; 2001US-0329414P.
PR	17-OCT-2001; 2001US-0330142P.
PR	18-OCT-2001; 2001US-0330309P.
PR	22-OCT-2001; 2001US-0341058P.
PR	24-OCT-2001; 2001US-0332666P.
PR	24-OCT-2001; 2001US-0343629P.
PR	29-OCT-2001; 2001US-0349575P.
PR	01-NOV-2001; 2001US-0346357P.
PR	17-APR-2002; 2002US-0373260P.
PR	19-APR-2002; 2002US-0373815P.
PR	19-APR-2002; 2002US-0373817P.
PR	19-APR-2002; 2002US-0373826P.
PR	19-APR-2002; 2002US-0373884P.
PR	22-APR-2002; 2002US-0374977P.
PR	16-MAY-2002; 2002US-0381037P.
PR	16-MAY-2002; 2002US-0381038P.
PR	16-MAY-2002; 2002US-0381042P.
PR	17-MAY-2002; 2002US-0381642P.
PR	28-MAY-2002; 2002US-0383656P.
PR	29-MAY-2002; 2002US-0383811P.
PR	25-JUN-2002; 2002US-0391335P.
PR	01-OCT-2002; 2002US-00262511.
XX	
XX	(CURA-) CURAGEN CORP.
PI	Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI	Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI	Ort T, Gorlan K, Zehnusen BD, Anderson DW, Zhong M, Catterton E;
PI	Ji W, Miller CE, Rastelli L, Stone DJ, Pena CBA, Shenoy SG;
PI	Shinkets RA, Rotherberg ME, Leach WD, Agee M, Berghs C, Dipippo VA;
PI	Eisen AU, Gangolli EA, Rieger DK, Spaderna SK;
XX	
DR	WPI; 2003-381626/36.
DR	N-PSDB; ADA05653.
XX	
PT	New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT	preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT	cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT	pharmacogenomics.
XX	
XX	Claim 1; Page 105-106; 586pp; English.
XX	
CC	The present invention describes NOVX proteins, where X can be 1 to 55
CC	(e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC	described above and a carrier; (2) a kit comprising, in one or more
CC	containers, the composition described above; (3) an isolated nucleic acid
CC	molecule which encodes a NOVX protein of the invention; (4) a vector
CC	comprising the nucleic acid molecule described above; (5) a cell
CC	comprising the above vector; (6) an antibody that immunospecifically
CC	binds to the polypeptide described above; (7) methods for determining the
CC	presence or amount of the above polypeptide or nucleic acid molecule in a
CC	sample; (8) methods for determining the presence of or predisposition to a
CC	disease associated with altered levels of expression of the above
CC	polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC	method of identifying an agent that binds to the polypeptide described

CC	above; (10) a method for identifying a potential therapeutic agent for
CC	use in treating a pathology that is related to an aberrant expression or
CC	aberrant physiological interactions of the polypeptide; (11) a method of
CC	screening for a modulator of activity or of latency or predisposition to
CC	a pathology associated with the polypeptide; (12) a method for modulating
CC	the activity of the polypeptide described above; (13) methods of treating
CC	or preventing a pathology associated with the above polypeptide in a
CC	mammal; and (14) a method for producing the above polypeptide. NOVX
CC	sequences have antidiabetic, anorectic, antibacterial, virucide, and
CC	immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
CC	and antilipase activities, and can be used in gene therapy. The
CC	polypeptide is useful in manufacturing a medicament for treating a
CC	syndrome associated with a human disease. The polypeptide or the nucleic
CC	acid molecule may be used to diagnose, treat or prevent metabolic
CC	disorders such as diabetes or obesity, infections, cachexia, cancer,
CC	neurodegenerative disorders, such as Alzheimer's disease or Parkinson's
CC	disease, immune disorders, haematopoietic disorders and various
CC	dyslipidaemias. The nucleic acids can also be used as hybridisation
CC	probes, in chromosome mapping, tissue typing, preventive medicine and
CC	pharmacogenomics. The present sequence represents a human NOVX from the
CC	present invention.
XX	
XX	Sequence 2066 AA;..
XX	
QY	Query Match 95.4%; Score 10007; DB 6; Length 2066;
QY	Best Local Similarity 95.6%; Pred. No. 0;
QY	Matches 1979; Conservative 3; Mismatches 7; Indels 82; Gaps 4;
QY	
QY	1 MLKFKYCARNPLDAGAAEPIASRASRLNLFQCKPPEWTCQMSPLSREGILDALFVLFE 60
QY	1 MLKFKYCARNPLDAGAAEPIASRASRLNLFQCKPPEWTCQMSPLSREGILDALFVLFE 60
QY	61 ECSQPALMKIKHVSNFVRKYSNTIAELQELQPSAKDFEVRSLVGCCHFAEVVVREKATG 120
QY	61 ECSQPALMKIKHVSNFVRKYSNTIAELQELQPSAKDFEVRSLVGCCHFAEVVVREKATG 120
QY	121 DIYAMKMKKALLAQOVSFEEERNILSRSTSPWIPOLQYAFQDKXNHLVMEYPPGG 180
QY	121 DIYAMKMKKALLAQOVSFEEERNILSRSTSPWIPOLQYAFQDKXNHLVMEYPPGG 180
QY	181 DLSLLNRYEDQDENLIQFYLAELIIVAVSHVMGVVHRDIPENILVDRTHIKLVDF 240
QY	181 DLSLLNRYEDQDENLIQFYLAELIIVAVSHVMGVVHRDIPENILVDRTHIKLVDF 240
QY	241 GSAAKMSNKMVNAKLPIGTDPYMAPVLTVMGDGGKTYGLDCDMWSVGVIAEMTYGR 300
QY	241 GSAAKMSNKMVNAKLPIGTDPYMAPVLTVMGDGGKTYGLDCDMWSVGVIAEMTYGR 300
QY	301 SPRAEGTSARTFNIMNFQRLKPPDDPKYSSDFLDLIQSLCGQKRLKFEGLCCCHPEF 360
QY	301 SPRAEGTSARTFNIMNFQRLKPPDDPKYSSDFLDLIQSLCGQKRLKFEGLCCCHPEF 360
QY	361 SKIDWNNIRNSPPFPVPTLTKSDDDTSNFDPEKNSWSSSPCQLSPSGFSGEELFFVGF 420
QY	361 SKIDWNNIRNSPPFPVPTLTKSDDDTSNFDPEKNSWSSSPCQLSPSGFSGEELFFVGF 420
QY	421 YSKALGILGRSESVVSGLDSPAKTSMEKLLIKSKELQSDQKCHKQVEMTLHRVS 480
QY	421 YSKALGILGRSESVVSGLDSPAKTSMEKLLIKSKELQSDQKCHKQVEMTLHRVS 480
QY	481 EVEAVLSQKEVELKASQTSRLLEQDLATYITECSSLKRSLEQARMEVSQEDDKALQLLH 540
QY	481 EVEAVLSQKEVELKASQTSRLLEQDLATYITECSSLKRSLEQARMEVSQEDDKALQLLH 540
QY	541 DIRSQSKLQEIKEQYQAOVEENRMLMNOLEEDLVSAARRRDLVSESLRESRLAABEFK 600
QY	541 DIRSQSKLQEIKEQYQAOVEENRMLMNOLEEDLVSAARRRDLVSESLRESRLAABEFK 600
QY	601 RKATECHQKLLKADQCKPVEGAKLEKINAQQLKIQLQEKLEKAVKASATEATLQ 659
QY	601 RKATECHQKLLKADQCKPVEGAKLEKINAQQLKIQLQEKLEKAVKASATEATLQ 659
QY	649 ---AKERABRELEKLNQREDSSGIRKLVEAREERHSLNKKVRLTETMERENRLKDD 704

Db 660 NIRQAKERARELEKLNQRRDSEGRKLVLEAEERHSLNKKVLEWERNENLKKD 719  
Qy 705 IQTKSQIQQMAKILILEBKHEAQAQVSAQHLEVHLKQKQHYEYKIVLDNQIKKDLAD 764  
Db 720 IQTKSQIQQMAKILILEBKHEAQAQVSAQHLEVHLKQKQHYEYKIVLDNQIKKDLAD 779  
Qy 765 KETLENMQRHEBEAHEKGIILSEKAMINAMDSKIRSLFORIVELSEANKLAANSSLET 824  
Db 780 KETLENMQRHEBEAHEKGIILSEKAMINAMDSKIRSLFORIVELSEANKLAANSSLET 839  
Qy 825 QRNKAQEEIMISLRQKQFYLETQAGKLEAQNKLBEQLEKISHQDHSKNRILLELETRL 884  
Db 840 QRNKAQEEIMISLRQKQFYLETQAGKLEAQNKLBEQLEKISHQDHSKNRILLELETRL 899  
Qy 885 REVSLEHEBKLEIKROLTELQLSQERESQLTALQARAALLESQLRQKTELETTAEA 944  
Db 900 REVSLEHEBKLEIKROLTELQLSQERESQLTALQARAALLESQLRQKTELETTAEA 959  
Qy 945 EEBIQALTARHDIQKFDALRNSCTVITDLBEQLNQLTEDNAELNNQNFYLSKOLDEAS 1004  
Db 960 EEBIQALTARHDIQKFDALRNSCTVITDLBEQLNQLTEDNAELNNQNFYLSKOLDEAS 1019  
Qy 1005 GANDEIVQLRSEVDHLRREITEREMOLTSQKOTMEALKTTCTMLBEQVMDLBAALDELLE 1064  
Db 1020 GANDEIVQLRSEVDHLRREITEREMOLTSQKOTMEALKTTCTMLBEQVMDLBAALDELLE 1079  
Qy 1065 KERQWEAARVLDGKSOFCRVREIQRLMDTBKQSRARADQRIETESROVVELAVKEHKA 1124  
Db 1080 KERQWEAARVLDGKSOFCRVREIQRLMDTBKQSRARADQRIETESROVVELAVKEHKA 1139  
Qy 1125 EILALQALKEQKLKAPSLDKLNDLEKXHAMLENNARSLOQKLETERELKORLLEEQAK 1184  
Db 1140 EILALQALKEQKLKAPSLDKLNDLEKXHAMLENNARSLOQKLETERELKORLLEEQAK 1199  
Qy 1185 LOQOMDLQKXNIFRLTQGLQALDRADLLKTERSDILEYOLENIVLYSHEKVMGTTISQ 1244  
Db 1200 LOQOMDLQKXNIFRLTQGLQALDRADLLKTERSDILEYOLENIVLYSHEKVMGTTISQ 1259  
Qy 1245 QTKLIDFLQAKMDQPAKKKKGLFSRKEPDALPTQVPLQVNLKLALEKAKARCALEBA 1304  
Db 1260 QTKLIDFLQAKMDQPAKKK-----VPLQVNLKLALEKAKARCALEBA 1304  
Qy 1305 LQKTRIELRSAREAAHRKATDHPHSTATAPQQTAMSAIVSRPHQSAMSLAPPSS 1364  
Db 1305 LQKTRIELRSAREAAHRKATDHPHSTATAPQQTAMSAIVSRPHQSAMSLAPPSS 1364  
Qy 1365 RKESSTPEEFSRRLKERMHNNIPHPFNVLNNRATKCAVCLDTVHFGQASKCLEQVM 1424  
Db 1365 RKESSTPEEFSRRLKERMHNNIPHPFNVLNNRATKCAVCLDTVHFGQASKCLEQVM 1424  
Qy 1425 CHPKSTCPLPATCGLPAEYATHTEAFCDKXNNSPGLQKPESSSLHLEGWKKVPRNNKR 1484  
Db 1425 CHPKSTCPLPATCGLPAEYATHTEAFCDKXNNSPGLQKPESSSLHLEGWKKVPRNNKR 1484  
Qy 1485 GQGWDRKYIIVLGSKVLIYDNEAREAGORPVEEFELCLPDGDVSTHGAVGASELANTAK 1544  
Db 1485 GQGWDRKYIIVLGSKVLIYDNEAREAGORPVEEFELCLPDGDVSTHGAVGASELANTAK 1544  
Qy 1545 A-----PKABADAKL 1554  
Db 1545 ADVPYILKMESHPTTCWPGRTLYLLAPSPDKQRWVTALLESVWAGGRVREKAEADAKL 1604  
Qy 1555 LGNSLLKLGDDRLDNKNTLPSDOVVLVGTBEGLYALNVKNSLTHVPGICAVQIYII 1614  
Db 1605 LGNSLLKLGDDRLDNKNTLPSDOVVLVGTBEGLYALNVKNSLTHVPGICAVQIYII 1664  
Qy 1615 KDLEKLLMIAGEBRALCLVDVKKVQSLAQSHLPAQFPDISPNIFFAVKGCHLFGAGKIEN 1674  
Db 1665 KDLEKLLMIAGEBRALCLVDVKKVQSLAQSHLPAQFPDISPNIFFAVKGCHLFGAGKIEN 1724  
Qy 1675 GLCICAAKPSKVILRYNENLSKYCIKREIETSEPCSCIHTFNYSILIGTNKFYEIDMKQ 1734

Db 1725 GLCICAAKPSKVILRYNENLSKYCIKREIETSEPCSCIHTFNYSILIGTNKFYEIDMKQ 1784  
Qy 1735 YTLBEFLDKNDHSLAPAVFAAASNSFPVSVIVQNSAGQREBYLLCFHFGVFDVDSYGRRS 1794  
Db 1785 YTLBEFLDKNDHSLAPAVFAAASNSFPVSVIVQNSAGQREBYLLCFHFGVFDVDSYGRRS 1844  
Qy 1795 RTDDLKWSRLPLAFAYREPVLVFTHFNLSLEVIEIQARSSAGTTPARAVLDIPNRYLGPAL 1854  
Db 1845 RTDDLKWSRLPLAFAYREPVLVFTHFNLSLEVIEIQARSSAGTTPARAVLDIPNRYLGPAL 1904  
Qy 1855 SSGAIYLASSYQDKLRLVICCKNLVKESGTEHHRGPTSSRSPNKRGPPTTYNEHITKRV 1914  
Db 1905 SSGAIYLASSYQDKLRLVICCKNLVKESGTEHHRGPTSSRSPNKRGPPTTYNEHITKRV 1964  
Qy 1915 SSPAPPGSPHPRPSTPHRYREGRTLRDKSPGRPLEREKSPGRMLSTERSSPGRLLF 1974  
Db 1965 SSPAPPGSPHPRPSTPHRYREGRTLRDKSPGRPLEREKSPGRMLSTERSSPGRLLF 2024  
Qy 1975 EDSRGRIPAGAVRTPLSQVKNKRGQSASQV 2005  
Db 2025 EDSRGRIPAGAVRTPLSQVKNKRGQSASQV 2055  
RESULT 8  
ABG78363  
ID ABG78363 standard; protein; 2053 AA.  
XX  
AC ABG78363;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE RHO/RAC-interacting citron kinase-like human protein, designated NOV3b.  
XX  
KW Human; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;  
KW atherosclerosis; diabetes; cell signalling; metabolic pathway;  
KW cellular receptor; downstream effector; cancer; gene therapy;  
KW hypertension; congenital heart defect; aortic stenosis; obesity;  
KW infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;  
KW neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;  
KW haematopoietic disease; scleroderma; fertility; immunogen;  
KW idiopathic thrombocytopenic purpura; graft versus host disease;  
KW Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;  
KW systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;  
KW stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;  
KW pain; alcoholism; transgenic.  
OS Homo sapiens.  
XX  
XX WO200226826-A2.  
XX  
XX 04-APR-2002.  
XX  
XX 27-SEP-2001; 2001WO-US042336.  
XX  
XX 27-SEP-2000; 2000US-0235631P.  
PR 27-SEP-2000; 2000US-0235631P.  
PR 27-SEP-2000; 2000US-0235631P.  
PR 27-SEP-2000; 2000US-0235808P.  
PR 27-SEP-2000; 2000US-0236064P.  
PR 27-SEP-2000; 2000US-0236065P.  
PR 27-SEP-2000; 2000US-0236065P.  
PR 28-SEP-2000; 2000US-0236135P.  
PR 03-OCT-2000; 2000US-0237434P.  
PR 05-OCT-2000; 2000US-0238321P.  
PR 06-OCT-2000; 2000US-0238356P.  
PR 06-OCT-2000; 2000US-0238399P.  
PR 16-MAR-2001; 2001US-0276667P.  
PR 31-MAY-2001; 2001US-0294823P.  
PR 12-JUL-2001; 2001US-0304868P.  
PR 26-SEP-2001; 2001US-00235631.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Gerlach VL, Macdougall JR, Smithson G, Millet I, Stone D;

PI Gunther E, Ellerman K, Grosse WM, Alsobrook JP, Lopley DW;  
PT Burgess CE, Radigaru M, Kekuda R, Spytek KA, Leach MD, Shinkets RA;  
XX WPI: 2002-499860/53.  
DR N-PSDB; ABS63436.  
XX  
XX Novel isolated NOVX polypeptides and polynucleotides homologous to  
PT attractin, plexin, papin-like family of proteins, useful for treating  
PT atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and  
PT stroke.  
XX  
PS Claim 1; Page 44-45; 309pp; English.  
XX  
CC The invention discloses the isolated human polypeptides, and  
CC polynucleotides encoding them, that have been designated NOVX. The  
CC polypeptides, polynucleotides and antibodies are useful in treating or  
CC preventing a NOVX-associated disorder which is cardiomyopathy, a  
CC atherosclerosis and diabetes in a human, where the disorder is related to  
CC cell signal processing and metabolic pathway modulation. They can also be  
CC used in determining the presence of, or predisposition to, a disease  
CC associated with altered levels of the polypeptides and polynucleotides of  
CC any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for  
CC identifying an agent that binds to, or that modulates the expression or  
CC activity of the polypeptide, for identifying an agent which is cellular  
CC receptor or downstream effector, for treating or preventing a NOVX-  
CC associated disorder and as a pharmaceutical composition comprising the  
CC polypeptide, polynucleotide or the antibody. The polypeptides and  
CC polynucleotides are useful in diagnostic applications (e.g. as a marker  
CC for cancerous cells or tissue types) where their amounts are assessed, or  
CC for the manufacture of a medicament (e.g. gene therapy) for treating or  
CC preventing disorders or syndromes such as hypertension, congenital heart  
CC defects, aortic stenosis, obesity, infectious disease, neurodegenerative disorders,  
CC Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,  
CC haemophilia, dyslipidemias, haematopoietic diseases, scleroderma,  
CC fertility, idiopathic thrombocytopenic purpura, graft versus host  
CC diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune  
CC disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,  
CC allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar  
CC ataxia, pain and alcoholism. They may also be used as immunogens to  
CC produce antibodies specific for the invention, and as vaccines.  
CC Transgenic cells containing a NOVX expressing construct are useful to  
CC produce non-human transgenic animals for studying the function and/or  
CC activity of the NOVX proteins and for identifying and/or evaluating  
CC modulators of NOVX protein activity. Transgenic cells containing a NOVX  
CC expressing construct are useful to produce non-human transgenic animals  
CC for studying the function and/or activity of the NOVX proteins and for  
CC identifying and/or evaluating modulators of NOVX protein activity. The  
CC sequences presented in ABG78359-ABG78371 are the human NOV1-NOV8 proteins  
XX  
XX Sequence 2053 AA;  
XX  
Query Match 95.4%; Score 10005; DB 5; Length 2053;  
Best Local Similarity 95.6%; Pred. No. 0;  
Matches 1978; Conservative 3; Mismatches 5; Indels 82; Gaps 4;  
QY 1 MLKFKYGARNPLDAGAAEPIAGRASRLNLFPGKPPFMTQQQWSPLSREGILDALFVLFE 60  
DB 1 MLKFKYGARNPLDAGAAEPIAGRASRLNLFPGKPPFMTQQQWSPLSREGILDALFVLFE 60  
QY 61 ECSQPALMKIKHVSFNVRKYSTTIAELQELQPSAKDFEVRSLVGCGHFAEVQVREKATG 120  
DB 61 ECSQPALMKIKHVSFNVRKYSTTIAELQELQPSAKDFEVRSLVGCGHFAEVQVREKATG 120  
QY 121 DIYAMKVMKKKALLAQEQVSFEERNILSRSTSPWIPQLOAFQDKNHLYLWNEYQPGG 180  
DB 121 DIYAMKVMKKKALLAQEQVSFEERNILSRSTSPWIPQLOAFQDKNHLYLWNEYQPGG 180  
QY 181 DLLSLNRYEDQDENLIFYLAELLIAVSHVLMGYVHRDKPENILVDRGTGHIKLVDVF 240  
DB 181 DLLSLNRYEDQDENLIFYLAELLIAVSHVLMGYVHRDKPENILVDRGTGHIKLVDVF 240  
QY 241 GSAAKVNSKNNVNAKPIGTPDYMAPEVLTVMGDGGKGTGYGLDCDWSVGVYAIYEMIYGR 300  
DB 241 GSAAKVNSKNNVNAKPIGTPDYMAPEVLTVMGDGGKGTGYGLDCDWSVGVYAIYEMIYGR 300

DB 241 GSAAKVNSKNNVNAKPIGTPDYMAPEVLTVMGDGGKGTGYGLDCDWSVGVYAIYEMIYGR 299  
QY 301 SPRAEGTSARTENINNFQFLKFPDDPKVSSDFLDLIQSILCGQKRLKFEGLCCCHPFF 360  
DB 300 SPFAEGTSARTFNIMNFQFLKFPDDPKVSSDFLDLIQSILCGQKRLKFEGLCCCHPFF 359  
QY 361 SKIDWNINRSPPPVPTLKSDDDTSNFDEPEKNSWVSSPCQLSPSGFSGEELPFVGF 420  
DB 360 SKIDWNINRSPPPVPTLKSDDDTSNFDEPEKNSWVSSPCQLSPSGFSGEELPFVGF 419  
QY 421 YSKALGILGRSESVVGLDSPAKTSSMEKLLIKSKELQSDQKCHKEQEMTRLHRVS 480  
DB 420 YSKALGILGRSESVVGLDSPAKTSSMEKLLIKSKELQSDQKCHKEQEMTRLHRVS 479  
QY 481 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSLKRSLEQARMVSEDOKALQLH 540  
DB 480 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSSLKRSLEQARMVSEDOKALQLH 539  
QY 541 DIREQSKLQEIKEQEQVQAQVEENRIMNQLEEDLVSAARRSDLYESELRESLAAEFK 600  
DB 540 DIREQSKLQEIKEQEQVQAQVEENRIMNQLEEDLVSAARRSDLYESELRESLAAEFK 599  
QY 601 RKATECQKLLKAKDQKPEVGEVAKLEKINAEQOLKIQLQEKLEKA----- 648  
DB 600 RKATECQKLLKAKDQKPEVGEVAKLEKINAEQOLKIQLQEKLEKA----- 659  
QY 649 ----AKERAERELEKLNREDSSEGIKKLVEABERHSLKVKRLTETMERENRLKDD 704  
DB 660 NIROKERAERELEKLNREDSSEGIKKLVEABERHSLKVKRLTETMERENRLKDD 719  
QY 705 IQTSQOIQOMADKILEBKHEQAQVSAOHLVHLKQKQHYEETKVLQDNQKKDLAD 764  
DB 720 IQTSQOIQOMADKILEBKHEQAQVSAOHLVHLKQKQHYEETKVLQDNQKKDLAD 779  
QY 765 KETLENMQRHEEBAHEKGIKLSQKAMINAMDSKIRSLQRIVELSEANKLAANSILFT 824  
DB 780 KETLENMQRHEEBAHEKGIKLSQKAMINAMDSKIRSLQRIVELSEANKLAANSILFT 839  
QY 825 QRNKAQEBEMISELRQKQFYLETOAGKLEAQNRLKEQLEKISHQDSDKNRLLETRL 884  
DB 840 QRNKAQEBEMISELRQKQFYLETOAGKLEAQNRLKEQLEKISHQDSDKNRLLETRL 899  
QY 885 REVSLHEEOKLEKROLTELQLSQERESQTLQAARAALLESQLEAETFEETAAE 944  
DB 900 REVSLHEEOKLEKROLTELQLSQERESQTLQAARAALLESQLEAETFEETAAE 959  
QY 945 EEEIQAULTAHRDEIQRFDAIRNSCTVITDLEBOLNQTEDNAELNNQNFYLSKQLEAS 1004  
DB 960 EEEIQAULTAHRDEIQRFDAIRNSCTVITDLEBOLNQTEDNAELNNQNFYLSKQLEAS 1019  
QY 1005 GANDEIVQLRSEVDHLREITEREMQLTSSOKTMEALKTCTMLEBQVMDLEALNDELLE 1064  
DB 1020 GANDEIVQLRSEVDHLREITEREMQLTSSOKTMEALKTCTMLEBQVMDLEALNDELLE 1079  
QY 1065 KERQWEAWRSVLGDEKSFQFECRVRELQRMQDTEKQSRARADQRTESQVVELAVKEHKA 1124  
DB 1080 KERQWEAWRSVLGDEKSFQFECRVRELQRMQDTEKQSRARADQRTESQVVELAVKEHKA 1139  
QY 1125 EILAQALKEQKLEKSLDKLNDLEKHAMLENNARSLQOQLETERELKQRLLEQAK 1184  
DB 1140 EILAQALKEQKLEKSLDKLNDLEKHAMLENNARSLQOQLETERELKQRLLEQAK 1199  
QY 1185 LQQQMDLQKNHIFRLTQGLQBALDRADLLKTERSDLEYQLENQVLYSHEKVKMEGTSIQ 1244  
DB 1200 LQQQMDLQKNHIFRLTQGLQBALDRADLLKTERSDLEYQLENQVLYSHEKVKMEGTSIQ 1259  
QY 1245 QTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQINELKLALEKSKARCAELEEA 1304  
DB 1260 QTKLIDFLQAKMDQPAKKK-----VPLQINELKLALEKSKARCAELEEA 1304  
QY 1305 LQKTRIELRSAREEAAHRKATDHPSTPATARQOIAMSAIVRSPHQPAMSLLAPSS 1364  
DB 1305 LQKTRIELRSAREEAAHRKATDHPSTPATARQOIAMSAIVRSPHQPAMSLLAPSS 1364

Qy	1365	RRKESSTPEEFSSRLKERMHNH:PHRFNVGLNMRATKCAVCLDTHVHFGQASKLECCQVM	1424	PF	02-OCT-2002;	2002WO-US031373.
Db	1365	RRKESSTPEEFSSRLKERMHNH:PHRFNVGLNMRATKCAVCLDTHVHFGQASKLECCQVM	1424	XX	02-OCT-2001;	2001US-0326483P.
Qy	1425	CHPKCSTCLPATCGLPAYATHTEAFCDKMNKSPGLQKEPSSSLHLEGWKKVPRNNK	1484	PR	05-OCT-2001;	2001US-0327435P.
Db	1425	CHPKCSTCLPATCGLPAYATHTEAFCDKMNKSPGLQKEPSSSLHLEGWKKVPRNNK	1484	PR	05-OCT-2001;	2001US-0327449P.
Qy	1485	GOQGWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAK	1544	PR	09-OCT-2001;	2001US-0327917P.
Db	1485	GOQGWDRKYIVLEGSKVLIYDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANTAK	1544	PR	09-OCT-2001;	2001US-0328029P.
Qy	1545	A-----EKAADAKL	1554	PR	09-OCT-2001;	2001US-0328044P.
Db	1545	ADVPYILKMESHPTTCWPQRTLYLLAPSPDQQRWVTALESVVAGGRVSRKAEADAKL	1604	PR	09-OCT-2001;	2001US-0328056P.
Qy	1555	LGNLSLLKLEGGDRLDNMCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYII	1614	PR	12-OCT-2001;	2001US-0328849P.
Db	1555	LGNLSLLKLEGGDRLDNMCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYII	1614	PR	12-OCT-2001;	2001US-0328856P.
Qy	1615	KDLEKLMIAGERALCLVDVKKVQSLAQSHLPAQPDISPNIFFAVKGCFLFGAGKIKEN	1674	PR	15-OCT-2001;	2001US-0329414P.
Db	1615	KDLEKLMIAGERALCLVDVKKVQSLAQSHLPAQPDISPNIFFAVKGCFLFGAGKIKEN	1674	PR	15-OCT-2001;	2001US-0330142P.
Qy	1675	GLCICAAAMPKSVILRYNENLSKYCIKETESEPCSCITHFTNYSILIGTNKPFYEDMKQ	1734	PR	17-OCT-2001;	2001US-0330309P.
Db	1675	GLCICAAAMPKSVILRYNENLSKYCIKETESEPCSCITHFTNYSILIGTNKPFYEDMKQ	1734	PR	17-OCT-2001;	2001US-0341058P.
Qy	1725	GLCICAAAMPKSVILRYNENLSKYCIKETESEPCSCITHFTNYSILIGTNKPFYEDMKQ	1784	PR	22-OCT-2001;	2001US-0339266P.
Db	1725	GLCICAAAMPKSVILRYNENLSKYCIKETESEPCSCITHFTNYSILIGTNKPFYEDMKQ	1784	PR	22-OCT-2001;	2001US-0339266P.
Qy	1735	YTLLEFLDKNDHSLAPAVFAASSNFPVSIQVNSAGQREYLLCFFHEFGVFDVSYGRS	1794	PR	24-OCT-2001;	2001US-0343629P.
Db	1735	YTLLEFLDKNDHSLAPAVFAASSNFPVSIQVNSAGQREYLLCFFHEFGVFDVSYGRS	1794	PR	24-OCT-2001;	2001US-0343629P.
Qy	1785	YTLLEFLDKNDHSLAPAVFAASSNFPVSIQVNSAGQREYLLCFFHEFGVFDVSYGRS	1844	PR	29-OCT-2001;	2001US-0349575P.
Db	1785	YTLLEFLDKNDHSLAPAVFAASSNFPVSIQVNSAGQREYLLCFFHEFGVFDVSYGRS	1844	PR	29-OCT-2001;	2001US-0349575P.
Qy	1795	RTDDLKWSRLPLAFAYREPFLVTHFNSLEVIQIARSSAGTPARAYLDIPNRYLGPAL	1854	PR	01-NOV-2001;	2001US-0346357P.
Db	1795	RTDDLKWSRLPLAFAYREPFLVTHFNSLEVIQIARSSAGTPARAYLDIPNRYLGPAL	1854	PR	01-NOV-2001;	2001US-0346357P.
Qy	1845	RTDDLKWSRLPLAFAYREPFLVTHFNSLEVIQIARSSAGTPARAYLDIPNRYLGPAL	1904	PR	17-APR-2002;	2002US-0373260P.
Db	1845	RTDDLKWSRLPLAFAYREPFLVTHFNSLEVIQIARSSAGTPARAYLDIPNRYLGPAL	1904	PR	17-APR-2002;	2002US-0373260P.
Qy	1855	SSGAIVLASSYQDKLAVICCKGNLVKESGTEHHRGPGSTSRSPNKGPGPTYNEHITKVA	1914	PR	16-MAY-2002;	2002US-0381037P.
Db	1855	SSGAIVLASSYQDKLAVICCKGNLVKESGTEHHRGPGSTSRSPNKGPGPTYNEHITKVA	1914	PR	16-MAY-2002;	2002US-0381037P.
Qy	1905	SSGAIVLASSYQDKLAVICCKGNLVKESGTEHHRGPGSTSRSPNKGPGPTYNEHITKVA	1964	PR	16-MAY-2002;	2002US-0381042P.
Db	1905	SSGAIVLASSYQDKLAVICCKGNLVKESGTEHHRGPGSTSRSPNKGPGPTYNEHITKVA	1964	PR	16-MAY-2002;	2002US-0381042P.
Qy	1915	SSPAPPEGSHPREPSTPHRYREGTELRDKSPGRLPERSKPGRLSTRERSPGRLUF	1974	PR	17-MAY-2002;	2002US-0381642P.
Db	1915	SSPAPPEGSHPREPSTPHRYREGTELRDKSPGRLPERSKPGRLSTRERSPGRLUF	1974	PR	17-MAY-2002;	2002US-0381642P.
Qy	1965	SSPAPPEGSHPREPSTPHRYREGTELRDKSPGRLPERSKPGRLSTRERSPGRLUF	2024	PR	28-MAY-2002;	2002US-0383656P.
Db	1965	SSPAPPEGSHPREPSTPHRYREGTELRDKSPGRLPERSKPGRLSTRERSPGRLUF	2024	PR	28-MAY-2002;	2002US-0383656P.
Qy	1975	EDSSRGLPAGAVRTPLSQVKNKGOSA	2002	PR	29-MAY-2002;	2002US-0383831P.
Db	1975	EDSSRGLPAGAVRTPLSQVKNKGOSA	2002	PR	29-MAY-2002;	2002US-0383831P.
Qy	2025	EDSSRGLPAGAVRTPLSQVKNKWDQSS	2052	PR	25-JUN-2002;	2002US-0391335P.
Db	2025	EDSSRGLPAGAVRTPLSQVKNKWDQSS	2052	PR	25-JUN-2002;	2002US-0391335P.
RESULT 9				XX	01-OCT-2002;	2002US-00262511.
ADA05642				PA	(CURA-) CURAGEN CORP.	
ID	ADA05642	standard; protein; 2053 AA.		XX	Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;	
XX				PI	Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;	
AC	ADA05642;			PI	Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;	
XX				PI	Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;	
DT	06-NOV-2003 (first entry)			PI	Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;	
XX				PI	Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;	
DE	Human NOVIa protein SEQ ID NO:2.			XX	WPI; 2003-381626/36.	
XX				DR	N-PSDB; ADA05641.	
KW	human; NOVIa; antidiabetic; anorectic; antibacterial; virucide;			PT	New NOVIa polypeptides and nucleic acids, useful for diagnosing,	
KW	immunomodulator; cytostatic; nootropic; neuroprotective;			PT	preventing or treating NOVIa-associated disorders, e.g. diabetes, obesity,	
KW	antiparkinsonian; antidiabetic; gene therapy; human disease;			PT	cancer or dyslipidemia, and in chromosome mapping, tissue typing or	
KW	metabolic disorder; diabetes; obesity; infection; cachexia; cancer;			PT	pharmacogenomics.	
KW	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;			PS	Claim 1; Page 99-100; 586pp; English.	
XX	immune disorder; haematopoietic disorder; dyslipidaemia.			CC	The present invention describes NOVIa proteins, where X can be 1 to 55	
XX	Homo sapiens.			CC	(e.g. NOVIa). Also described: (1) a composition comprising a polypeptide	
XX	WO2003029424-A2.			CC	described above and a carrier; (2) a kit comprising, in one or more	
XX	10-APR-2003.			CC	containers, the composition described above; (3) an isolated nucleic acid	
XX				CC	molecule which encodes a NOVIa protein of the invention; (4) a vector	
XX				CC	comprising the nucleic acid molecule described above; (5) a cell	
XX				CC	comprising the above vector; (6) an antibody that immunospecifically	
XX				CC	binds to the polypeptide described above; (7) methods for determining the	
XX				CC	presence or amount of the above polypeptide or nucleic acid molecule in a	
XX				CC	sample; (8) methods for determining the presence of or predisposition to	
XX				CC	a disease associated with altered levels of expression of the above	
XX				CC	polypeptide or nucleic acid molecule in a first mammalian subject; (9) a	
XX				CC	method of identifying an agent that binds to the polypeptide described	
XX				CC	above; (10) a method for identifying a potential therapeutic agent for	
XX				CC	use in treating a pathology that is related to an aberrant expression or	
XX				CC	aberrant physiological interactions of the polypeptide; (11) a method of	
XX				CC	screening for a modulator of activity or of latency or predisposition to	
XX				CC	a pathology associated with the polypeptide; (12) a method for modulating	
XX				CC	the activity of the polypeptide described above; (13) methods of treating	
XX				CC	or preventing a pathology associated with the above polypeptide in a	
XX				CC	mammal; and (14) a method for producing the above polypeptide. NOVIa	

CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
CC and antipalemic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
CC present invention.  
XX  
SQ Sequence 2053 AA;

Query Match 95.4%; Score 10005; DB 6; Length 2053;  
Best Local Similarity 95.6%; Pred. No. 0;  
Matches 1978; Conservative 3; Mismatches 5; Indels 82; Gaps 4;  
QY 1 MLKFKYGASNPILDAGAAEPIASRASLNLFPGQKPPFMTQQQMSPLSRBGLDALFVLFE 60  
DB 1 MLKFKYGASNPILDAGAAEPIASRASLNLFPGQKPPFMTQQQMSPLSRBGLDALFVLFE 60  
QY 61 ECSQPALMKIKHVSNFVRKYSITIAELQELQPSAKDFEVRSLVGGCHFAEVQVREKATG 120  
DB 61 ECSQPALMKIKHVSNFVRKCSDTIAELQELQPSAKDFEVRSLVGGCHFAEVQVREKATG 120  
QY 121 DIYAMKMKKALLAQEQVSPFEBERNILSRSTSPWIPOLQYAFQDKHLYLMESYQFGG 180  
DB 121 DIYAMKMKKALLAQEQVSPFEBERNILSRSTSPWIPOLQYAFQDKHLYLMESYQFGG 180  
QY 181 DLLSLNRYEDQDENLIFYLAELILAVSHVLMGYVHRDIKPNILVDRTGHIKLVD 240  
DB 181 DLLSLNRYEDQDENLIFYLAELILAVSHVLMGYVHRDIKPNILVDRTGHIKLVD 240  
QY 241 GSAAKNSNMVNAKLPICPTDYMAPEVLTVNMNGDKGTGLDCDWSVGVYAYEMIYGR 300  
DB 241 GSAAKNSNK-VNAKLPICPTDYMAPEVLTVNMNGDKGTGLDCDWSVGVYAYEMIYGR 299  
QY 301 SPSAETSARTNNINNFRLKFPDDPKVSSDFDLTQSLLCGQKERLKEGLCCHPFF 360  
DB 300 SPSAETSARTNNINNFRLKFPDDPKVSSDFDLTQSLLCGQKERLKEGLCCHPFF 359  
QY 361 SKIDMNNIRNSPPFVPTLKSDDTSNFBEPKNSWSSSPQLSPGSPGSELPFGVGS 420  
DB 360 SKIDMNNIRNAPFVPTLKSDDTSNFBEPKNSWSSSPQLSPGSPGSELPFGVGS 419  
QY 421 YSKALCIGRSESVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHKMEQEMTRLHRRVS 480  
DB 420 YSKALCIGRSESVSGLDSPAKTSMEKKLLIKSKELQDSQDKCHKMEQEMTRLHRRVS 479  
QY 481 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSLKESLFOARMEVSQEDDKALQLLH 540  
DB 480 EVEAVLSQKEVELKASQTSRSLLEQDLATYITECSLKESLFOARMEVSQEDDKALQLLH 539  
QY 541 DIREQSRKLQETKEQSYQAVQEMRLMMNQLBEDLVSAARRSDLYSELSRSLAAAEFFK 600  
DB 540 DIREQSRKLQETKEQSYQAVQEMRLMMNQLBEDLVSAARRSDLYSELSRSLAAAEFFK 599  
QY 601 RKATECOHKLKAKQOQKPEVGEYAKLEKINAEQQLKIQELOEKLEKA----- 648  
DB 600 RKATECOHKLKAKQOQKPEVGEYAKLEKINAEQQLKIQELOEKLEKAVKASTATELLQ 659  
QY 649 ----AKERAERLEKLNQREDSEGIKKLVAEERRHSLNKVKELETMERENRLKDD 704  
DB 660 NTRQAKERAERLEKLNQREDSEGIKKLVAEERRHSLNKVKELETMERENRLKDD 719  
QY 705 IQTKSQIQOMADKILLEKREKREAOVSAQHLEVLKQEQHYBEKIKVLDNQIKDLAD 764  
DB 720 IQTKSQIQOMADKILLEKREKREAOVSAQHLEVLKQEQHYBEKIKVLDNQIKDLAD 779  
QY 765 KETLENMQRHBEAEKGIKLSQKAMINAMDSKIRSELQRIVELSEANKLAANSLSFT 824

DB 780 KETLENMQRHBEAEKGIKLSQKAMINAMDSKIRSELQRIVELSEANKLAANSLSFT 839  
QY 825 QRNKAQEMISELRQOKFYLETQAGKLEAQNRKLEQLEKISHQDSDKNRLLLETRL 884  
DB 840 QRNKAQEMISELRQOKFYLETQAGKLEAQNRKLEQLEKISHQDSDKNRLLLETRL 899  
QY 885 REVSLHBEQKLEKROLTQLQSLQRESQLTALQAARAALLESOLRQAKTELEETABEA 944  
DB 900 REVSLHBEQKLEKROLTQLQSLQRESQLTALQAARAALLESOLRQAKTELEETABEA 959  
QY 945 EEEIQALTARDEIQRFQDALRNSCTVITDLEQNLQTEDNNAELNNQNFYLSKQLEAS 1004  
DB 960 EEEIQALTARDEIQRFQDALRNSCTVITDLEQNLQTEDNNAELNNQNFYLSKQLEAS 1019  
QY 1005 GANDEIVOLRSEVDHLRREITEREMQLTISQQTNEALKTKTCTMLEBOVMOLEALNDELLE 1064  
DB 1020 GANDEIVOLRSEVDHLRREITEREMQLTISQQTNEALKTKTCTMLEBOVMOLEALNDELLE 1079  
QY 1065 KERQWEAWRSVLGDEKSFQFCRVRLEQRLMDTEKQSPARADQRTESQVQVLEAVKHKHA 1124  
DB 1080 KERQWEAWRSVLGDEKSFQFCRVRLEQRLMDTEKQSPARADQRTESQVQVLEAVKHKHA 1139  
QY 1125 EILALQALKEQKLEKASLSKDLNDEKHAMLEMNARSLOQKLETERELKQRLLEQAK 1184  
DB 1140 EILALQALKEQKLEKASLSKDLNDEKHAMLEMNARSLOQKLETERELKQRLLEQAK 1199  
QY 1185 LQQQMDLQKNHIFRLTQGLQELADRADLKTERTSDLEYQLENQVLYSHEKVKMEGTISQ 1244  
DB 1200 LQQQMDLQKNHIFRLTQGLQELADRADLKTERTSDLEYQLENQVLYSHEKVKMEGTISQ 1259  
QY 1245 QTKLIDFLQAKMDQPAKKKGLFSRRKEDPALPTQVPLQYNELKLALEKQKARCAELEEA 1304  
DB 1260 QTKLIDFLQAKMDQPAKKK-----VPLQYNELKLALEKQKARCAELEEA 1304  
QY 1305 LQKTRIELRSARBEAAHRKATDHPHPSTPATARQOQIAMSIVRSPHQPSAMSLLAPSS 1364  
DB 1305 LQKTRIELRSARBEAAHRKATDHPHPSTPATARQOQIAMSIVRSPHQPSAMSLLAPSS 1364  
QY 1365 RKESSTPEFSRRLKERMHNIPIHRENVGLNMRATKCAVCLDTVHFGROASKCLEQVQM 1424  
DB 1365 RKESSTPEFSRRLKERMHNIPIHRENVGLNMRATKCAVCLDTVHFGROASKCLEQVQM 1424  
QY 1425 CHPKXSTCLPATCGLPAEYATHFTEAFCRDKMNSPGLQTKPESSSLHLEGMKVPNRNKR 1484  
DB 1425 CHPKXSTCLPATCGLPAEYATHFTEAFCRDKMNSPGLQTKPESSSLHLEGMKVPNRNKR 1484  
QY 1485 GQQGMDRKYIVLEGSKVLIVDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASLANTAK 1544  
DB 1485 GQQGMDRKYIVLEGSKVLIVDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASLANTAK 1544  
QY 1545 A-----EKAEDAKL 1554  
DB 1545 ADVPYILKMESHPTTCWPGRTLYLLAPSPDKQWVTTALESVVAGGRVSRKAEADAKL 1604  
QY 1555 LGNSLLKLEGDRDLDMNCTLPFSDQVVLVGTGEGLYALNVKNSLTHVPQIGAVFYI 1614  
DB 1605 LGNSLLKLEGDRDLDMNCTLPFSDQVVLVGTGEGLYALNVKNSLTHVPQIGAVFYI 1664  
QY 1615 KDLEKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISFNIPFAVKGCHLFGAGKIEN 1674  
DB 1665 KOLEKLMIAAGEERALCLVDVKVKQSLAQSHLPAQPDISFNIPFAVKGCHLFGAGKIEN 1724  
QY 1675 GLCITCAAMPKSVILRYNENLSKYCIKTEITSEPCSHFTNYSILIGNKFYEIDMKQ 1734  
DB 1725 GLCITCAAMPKSVILRYNENLSKYCIKTEITSEPCSHFTNYSILIGNKFYEIDMKQ 1784  
QY 1735 YTLLEFLLDKNDHSLAPAVFAASNSPPVSIQVNSAGQREYLLCFHEFGFVDSYGRS 1794  
DB 1785 YTLLEFLLDKNDHSLAPAVFAASNSPPVSIQVNSAGQREYLLCFHEFGFVDSYGRS 1844  
QY 1795 RTDDLKWSRLPAPAYREPFLVTFHNSLEVIETIQARSSAGTTPARAYLDIPNRYLGPAL 1854

Db 1845 RTDCLKWSRLPLAFAYREPFLVTHFNLSLEVIEIQARSSAGTTPARAYLDIPNRYLGPAL 1904  
QY 1855 SSGAIVYASSYQDKLAVICCKGNLVKESGTEHHRGPSTSRSSPNKSGPPTYNEHITKRYA 1914  
Db 1905 SSGAIVYASSYQDKLAVICCKGNLVKESGTEHHRGPSTSRSSPNKSGPPTYNEHITKRYA 1964  
QY 1915 SSPAPPEGSPHPREPSTPHRYREGRTTELRRDKSPGRPLERKSPGRMLSTRERSRGLF 1974  
Db 1965 SSPAPPEGSPHPREPSTPHRYREGRTTELRRDKSPGRPLERKSPGRMLSTRERSRGLF 2024  
QY 1975 EDSRGRPLPAGAVRTPLSQVKNKRGQSA 2002  
Db 2025 EDSRGRPLPAGAVRTPLSQVKNKWDQSS 2052

RESULT 10  
ABP97683  
ID ABP97683 standard; protein; 2055 AA.  
XX  
AC ABP97683;  
XX  
DT 16-MAY-2003 (first entry)  
XX  
DE Polypeptide similar to citron rho/rac-interacting kinase-short kinase.  
XX  
KW Human; citron rho/rac-interacting kinase-short kinase; obesity;  
KW chronic obstructive pulmonary disease; hypertension; diabetes;  
KW coronary artery disease; hyperlipidaemia; stroke; gallbladder disease;  
KW gout; osteoarthritis; sleep apnea; cancer; thrombotic disease;  
KW polycystic ovarian syndrome; fertility; depression.  
XX  
OS Homo sapiens.  
XX  
PN W02003004629-A2.  
XX  
PD 16-JAN-2003.  
XX  
PF 01-JUL-2002; 2002WO-EP007229.  
XX  
PR 02-JUL-2001; 2001US-0301853P.  
PR 10-DEC-2001; 2001US-0337130P.  
PR 25-APR-2002; 2002US-0375015P.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Zhu Z;  
XX  
WP 2003-221595/21.  
XX  
DR New human citron rho/rac-interacting Kinase-short Kinase polypeptide and  
XX polynucleotide for preventing or treating diseases associated with the  
XX PT polypeptide dysfunction, e.g. obesity or chronic obstructive pulmonary  
XX PT disease.  
XX  
PS Disclosure; Fig 4; 145pp; English.  
XX  
CC The present sequence represents a polypeptide with some identity to a  
XX human citron rho/rac-interacting kinase-short kinase polypeptide. The  
XX CC polynucleotide and polypeptide of the invention are useful in preventing,  
XX CC ameliorating, or treating diseases associated with the polypeptide  
XX CC dysfunction. The expression vector or the reagent is useful in the  
XX CC preparation of a medicament for modulating the activity of a human citron  
XX CC rho/rac-interacting kinase-short kinase in a disease, such as obesity or  
XX CC chronic obstructive pulmonary disease. These may also be used for  
XX CC treating obesity/overweight-associated comorbidities, such as  
XX CC hypertension, diabetes, coronary artery disease, hyperlipidaemia, stroke,  
XX CC gallbladder disease, gout, osteoarthritis, sleep apnea, cancer,  
XX CC thrombotic diseases, polycystic ovarian syndrome, reduced fertility, and  
XX CC depression  
XX  
SQ Sequence 2055 AA;

Query Match

92.0%; Score 9656; DB 6; Length 2055;

Db 1020 GANDEIVQLRSEVDHLRRRITEREMQLTQSKQTMPEALKTCTTCLMEEQVLDLSEALNDELLE 1079  
QY 1065 KERQWEAWRSVLGDEKQSECEVRREIQRMLDTEKQSRARADQRIITESQVVELAVKEHKA 1124  
Db 1080 KERQWEAWRSVLGDEKQSECEVRREIQRMLDTEKQSRARADQRIITESQVVELAVKEHKA 1139  
QY 1125 EIALOQALKEOKLKAESLSDKLNLEKHAMLENNARSLOOKLETTERELKORLLEEQAK 1184  
Db 1140 EIALOQALKEOKLKAESLSDKLNLEKHAMLENNARSLOOKLETTERELKORLLEEQAK 1199  
QY 1185 LOQMDLQKNHIFRLTQGLQOALDRADLLKTERSDLEYOLENIQVLYSHEKVKMGTTISQ 1244  
Db 1200 LOQMDLQKNHIFRLTQGLQOALDRADLLKTERSDLEYOLENIQVLYSHEKVKMGTTISQ 1259  
QY 1245 QTKLIDFLOAKMDQPAKKKGLFSRKEDPALPTQVPLQYNELKLALEKEKARCAELEBEA 1304  
Db 1260 QTKLIDFLOAKMDQPAKKK-----VPLQYNELKLALEKEKARCAELEBEA 1304  
QY 1305 LOKTRIELSAREEAAHRAKATDHPHSTPATARQIAMSIVRSPEHQPMSALLAPSS 1364  
Db 1305 LOKTRIELSAREEAAHRAKATDHPHSTPATARQIAMSIVRSPEHQPMSALLAPSS 1364  
QY 1365 RKESSTPEFSGRLKERMHNI PHRFNVGLNMRATKCAVCLDTVHFGRAQSKLECCQM 1424  
Db 1365 RKESSTPEFSGRLKERMHNI PHRFNVGLNMRATKCAVCLDTVHFGRAQSKLECCQM 1424  
QY 1425 CHPKSTCLPATCGIPAEYATHFTEAFCDKXNSPGLQTKPEPSSSLHLBGWKKVPRNNKR 1484  
Db 1425 CHPKSTCLPATCGIPAEYATHFTEAFCDKXNSPGLQTKPEPSSSLHLBGWKKVPRNNKR 1484  
QY 1485 GQGGWDRKXIVLEGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1544  
Db 1485 GQGGWDRKXIVLEGSKVLIYDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANTAK 1544  
QY 1545 A-----EKABADAKL 1554  
Db 1545 ADVPYILKMHSHPTTCWPGRTLYLLAPSPDKQRMWTALESWAGVRVSRSEKABADAKL 1604  
QY 1555 LGNSILKLGDRDLNMCNLTLPESDQVVLVGTTEEGVALNVLKNSLTHVPGIGAVFQIVII 1614  
Db 1605 LGNSILKLGDRDLNMCNLTLPESDQVVLVGTTEEGVALNVLKNSLTHVPGIGAVFQIVII 1664  
QY 1615 KOLEKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDVSPNI FFAVKGCHLFGAGKIE 1674  
Db 1665 KOLEKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDVSPNI FFAVKGCHLFGAGKIE 1724  
QY 1675 GLCICAMPKSVVILRYNENLSKYCIKRIETSEPCSCIHFNYSILITGNTKFIIDMKQ 1734  
Db 1725 SLICICAMPKSVVILRYNENLSKYCIKRIETSEPCSCIHFNYSILITGNTKFIIDMKQ 1784  
QY 1735 YLLEFLDKNDHSLAPFAAASNSPPVSVIVQNSAGOREEYLLCFHFEGVFDVSYGRRS 1794  
Db 1785 YLLEFLDKNDHSLAPFAAASNSPPVSVIVQNSAGOREEYLLCFHFEGVFDVSYGRRS 1844  
QY 1795 RTDDLKWSRLPLAFAYREPYLFVTHNSLEVIETIOARSSAGTPAAYLDIPNRYLGPAT 1854  
Db 1845 RTDDLKWSRLPLAFAYREPYLFVTHNSLEVIETIOARSSAGTPAAYLEIPNRYLGPAT 1904  
QY 1855 SSGAIVLASSYQDKLRVICCKGNLVKESGTEHRRGPTSRSPKNGPPTTYNEHITKVA 1914  
Db 1905 SSGAIVLASSYQDKLRVICCKGNLVKESGTEHRRGPTSRSPKNGPPTTYNEHITKVA 1964  
QY 1915 SGPAPPEGSHPREPSTPHRY--REGRTLRDKSGPRLEREKSGPGRVLMSTRERSPCR 1972  
Db 1965 SGPAPPEGSHPREPSTPHRYDREGRTLRDKSGPRLEREKSGPGRVLMSTRERSPCR 2024  
QY 1973 LPESSRGRLPAGAVTPLSQVKNKGQSA 2002  
Db 2025 LPESSRGRLPAGAVTPLSQVKNKWQDQSS 2054

RESULT 11  
AA026960

AA026960 standard; protein; 2055 AA.  
AA026960;  
01-MAY-2003 (first entry)  
Human CR1K related protein sequence, SEQ ID No 3.  
Anorectic; hypotensive; cardiant; antilipaeic; cerebroprotective;  
anigout; osteopathic; antiarthritic; cytostatic; antidepressant;  
immunomodulator; antinamic; tranquiliser; antiparkinsonian; nootropic;  
neuroprotective; antinflammatory; antidiabetic; analgesic;  
human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;  
obesity; comorbidities; cancer; anorexia; cachexia; bulimia;  
central nervous system disorder; chronic obstructive pulmonary disease;  
diabetes; pain.  
Homo sapiens.  
WO2003004523-A1.  
16-JAN-2003.  
28-JUN-2002; 2002WO-EP007156.  
02-JUL-2001; 2001US-0301841P.  
11-DEC-2001; 2001US-0338651P.  
25-APR-2002; 2002US-0375014P.  
(FARB ) BAYER AG.  
Zhu Z;  
WPI; 2003-221576/21.  
New human citron rho/rac-interacting kinase (CR1K) polypeptide and  
polynucleotide, useful in preventing, ameliorating or treating diseases  
associated with human CR1K dysfunction, e.g. obesity, diabetes or  
Alzheimer's disease.  
Disclosure; Fig 3; 237pp; English.  
The invention relates to an isolated polynucleotide encoding a human  
citron rho/rac-interacting kinase polypeptide. The isolated  
polynucleotide comprises a 6165 or 8603 base pair sequence, given in the  
specification. The human citron rho/rac-interacting kinase (CR1K)  
polypeptide and polynucleotide are useful in preventing, ameliorating, or  
treating diseases associated with human CR1K dysfunction such as obesity  
and obesity-associated comorbidities (e.g. hypertension, coronary artery  
disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of  
cancer including endometrial, breast, prostate and colon cancer),  
anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood  
disorders, anxiety disorders, Parkinson's disease or Alzheimer's  
disease), chronic obstructive pulmonary disease, or diabetes. These can  
also be used to treat pain associated with the disorders. The human CR1K  
polypeptide is also useful in diagnostic assays or in genetic testing.  
The expression vector or the reagent is useful in preparing a medicament  
for modulating the activity of a human CR1K in a disease, e.g. obesity, a  
central nervous system disorder, or chronic obstructive pulmonary  
disease. The fusion protein is useful for generating antibodies against a  
CR1K polypeptide and for use in various assay systems. The methods are  
useful in producing and detecting the polynucleotide and polypeptide and  
in screening for agents that modulate the activity of the human CR1K  
polypeptide. This sequence represents a protein relating to the human  
CR1K protein of the invention  
Sequence 2055 AA;

Query Match 92.0%; Score 9656; DB 6; Length 2055;  
Best Local Similarity 91.9%; Pred. No. 0;  
Matches 1903; Conservative 41; Mismatches 42; Indels 84; Gaps 5;

QY 1 MLKFYGARNPDLGAAEPTASRASRLNLPFGQKPPFTQQXSPLSREGILDALFVLE 60

Db 1 MLAFKYGVRNPPASASEPTASASRLNFFQCKPPLMTCCQMSALSREGMLDALFALFE 60  
Qy 61 ECSQPALMKIKHYSNFKVYSDTIAELQELQPSAKOFVRSVLGCGHFAVQVVRREKATG 120  
Db 61 ECSQPALMKHVSFQKYSYSDTIAELRELQPSARDFEVRSVLGCGHFAVQVVRREKATG 120  
Qy 121 DIYAMKVMKKALLAQBOVSFFFEERNILSRSTS PMIPOLQYAFQDNHLYLMEEYQPGG 180  
Db 121 DVTAMKMKKALLAQBOVSFFFEERNILSRSTS PMIPOLQYAFQDNHLYLMEEYQPGG 180  
Qy 181 DLSSLNRYEQDQDENLIQIYLAELILAVHSVHMGVHRDIKPENILVDRTHIKLYDF 240  
Db 181 DFLSSLNRYEQDQDENLIQIYLAELILAVHSVHMGVHRDIKPENILVDRTHIKLYDF 240  
Qy 241 GSAAKMNSKMWNAKPIGTPDYMAPEVLTVMNGDCKGTGYGLDCDWKSVGVIAEYMIYGR 300  
Db 241 GSAAKMNSK- VDAKUPIGTPDYMAPEVLTVMNEDRRRTGYGLDCDWKSVGVIAEYMIYGR 299  
Qy 301 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDFDLIQSLGCGQKXERLKFGCLCHPFF 360  
Db 300 TPTEGTSARTFNNIMNFORFLKFPDDPKVSSBELLQSLGCGQKXERLKFGCLCHPFF 359  
Qy 361 SKIDWNIRNSPPFPVPTLKSDDDTSNFDEPEKNSVSSPCOLSPGSGSELFPVGES 420  
Db 360 ARTDWNIRNSPPFPVPTLKSDDDTSNFDEPEKNSWAFILCVPAEPFAFGSELFPVGES 419  
Qy 421 YSKALGILGRSESVSGLDSPAKTSMEKKLITKSKELQDSQDKHMQEMTRLHRRVS 480  
Db 420 YSKALGILGRSESVSGLDSPAKTSMEKKLITKSKELQDSQDKHMQEMTRLHRRVS 479  
Qy 481 EVBAVLSQKEVELKASQTSRSLLEQDLATYITPCSSLSKRSLEQARMEVSQEDDKALQLLH 540  
Db 480 EVBAVLSQKEVELKASQTSRSLLEQDLATYITPCSSLSKRSLEQARMEVSQEDDKALQLLH 539  
Qy 541 DIREQRKLOEIKBOQYQAOVEBRLMMQLEBDLVSAARRSDLYSELSRESRLAAEFK 600  
Db 540 DIREQRKLOEIKBOQYQAOVEBRLMMQLEBDLVSAARRSDLYSELSRESRLAAEFK 599  
Qy 601 RKATECOHKLKAKDQCKPEVGYAKLEKINAPQQLKIQELQEKLEKA 648  
Db 600 RKANECHKLKAKDQCKPEVGYSKLEKINAPQQLKIQELQEKLEKAVKASTEATELIQ 659  
Qy 649 ---- AKERABRELEKLNREDSSGTRKLVZAEERHSLENKYKLETMERENRLKDD 704  
Db 660 NIKQAKERAERLEKLNREDSSGTRKLVZAEERHSLENKYKLETMERENRLKDD 719  
Qy 705 IQTKSQIQQMAKILILEEKHREAQVSACHLEVHLKQEQHYEYEEKIKVLDNQIKKDLAD 764  
Db 720 IQTKSQIQQMAKILILEEKHREAQVSACHLEVHLKQEQHYEYEEKIKVLDNQIKKDLAD 779  
Qy 765 KETLENNQORHEEAEHKGKILSEQKAMINAMDSKIRSLBOQRIVELSEANKLAANSSUFT 824  
Db 780 KESLENNQORHEEAEHKGKILSEQKAMINAMDSKIRSLBOQRIVELSEANKLAANSSUFT 839  
Qy 825 QRNMKAQEMISLRQCKFYLETQACKLEBAQNKLEEQLEKISHQDHSKNLLELETRL 884  
Db 840 QRNMKAQEMISLRQCKFYLETQACKLEBAQNKLEEQLEKISHQDHSKNLLELETRL 899  
Qy 885 REVLSLHEBQKLEKRLQTELQLSQERESQLTALQARAALLESQLRQAKTELEETTAA 944  
Db 900 REVLSLHEBQKLEKRLQTELQLSQERESQLTALQARAALLESQLRQAKTELEETTAA 959  
Qy 945 EESIQALTARHDIQKFPALRNSCTVITDLBEQLNQLTEDNAELNNQNFYLSKOLDEAS 1004  
Db 960 EESIQALTARHDIQKFPALRNSCTVITDLBEQLNQLTEDNAELNNQNFYLSKOLDEAS 1019  
Qy 1005 GANDEIVQLRSEVDHLRRRITEREMOLTSQKQTMELAKTKTCTMVEQVMDLPAALDELLE 1064  
Db 1020 GANDEIVQLRSEVDHLRRRITEREMOLTSQKQTMELAKTKTCTMVEQVMDLPAALDELLE 1079  
Qy 1065 KERQWEAWSVLGDQKQPECRVRELQMLDTEKQSRABADQRTITESROVVELAVKEHKA 1124

Db 1080 KERQWEAWSVLGDQKQPECRVRELQMLDTEKQSRABADQRTITESROVVELAVKEHKA 1139  
Qy 1125 EILALQALKEQKLABSLDKLNDLEKKHAMLEMMNARSLOCKLETRELEKORLLEBOAK 1184  
Db 1140 EILALQALKEQKLABSLDKLNDLEKKHAMLEMMNARSLOCKLETRELEKORLLEBOAK 1199  
Qy 1185 LQQQMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYQLENIQVLYSHEKVMKEGTISQ 1244  
Db 1200 LQQQMDLQKNHIFRLTQGLQEAALDRADLLKTERSDLEYQLENIQVLYSHEKVMKEGTISQ 1259  
Qy 1245 QTKLIDFLQAKMDOPAKKXGLFSRRKEDPALPTQVPLQYNELKLALEKAKARCAELEEA 1304  
Db 1260 QTKLIDFLQAKMDOPAKKXGLFSRRKEDPALPTQVPLQYNELKLALEKAKARCAELEEA 1304  
Qy 1305 LOKTRIELRSAREAAHRKATDHPHPSTPATARQOIAMSAIVRSPEHQSAMSILLAPSS 1364  
Db 1305 LOKTRIELRSAREAAHRKATDHPHPSTPATARQOIAMSAIVRSPEHQSAMSILLAPSS 1364  
Qy 1365 RRKESSTPEFSRRLKERMHNIPHRNVGLNMRATKCAVCLDTVHFGROASKCLEQVM 1424  
Db 1365 RRKESSTPEFSRRLKERMHNIPHRNVGLNMRATKCAVCLDTVHFGROASKCLEQVM 1424  
Qy 1425 CHPKCSTCLPATCGLPAEYATHFTEAFCDKMSNPGLOTKEPSSSLHLEGWKMVPRNKR 1484  
Db 1425 CHPKCSTCLPATCGLPAEYATHFTEAFCDKMSNPGLOTKEPSSSLHLEGWKMVPRNKR 1484  
Qy 1485 GQQQWDRKYIVLEGSKVLIYDNEAREAGQRPVBEFELCLPDGDSVHGAAGSELANTAK 1544  
Db 1485 GQQQWDRKYIVLEGSKVLIYDNEAREAGQRPVBEFELCLPDGDSVHGAAGSELANTAK 1544  
Qy 1545 A-----EKAEADAKL 1554  
Db 1545 ADVPYILKMHESHPTTQWPGRTIYLLAPSPDKQWWTALLESVVGGRVREKAEADAKL 1604  
Qy 1555 LGNSLLKLEGDDRLDMNCTLPFSQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFQYII 1614  
Db 1605 LGNSLLKLEGDDRLDMNCTLPFSQVVLVGTBEGLYALNVLKNSLTHVPGIGAVFQYII 1664  
Qy 1615 KOLEKLMIAGEERALCLVDVKVKQSLAOSHUPAOPDISPNTFEAVKGCHELFAGKIE 1674  
Db 1665 KOLEKLMIAGEERALCLVDVKVKQSLAOSHUPAOPDISPNTFEAVKGCHELFAGKIE 1724  
Qy 1675 GLCICAAMPKSVTLRYNENLSKYCIKREIETSEPCSCIHFTNYSILIGTNKFEYIDMKQ 1734  
Db 1725 SLICICAAMPKSVTLRYNENLSKYCIKREIETSEPCSCIHFTNYSILIGTNKFEYIDMKQ 1784  
Qy 1735 YTLBEFLDKNDHSLAPAVFAASSNSFPVSVIQVNSAGQREYLLCPHEFGVFDVSYGRRS 1794  
Db 1785 YTLBEFLDKNDHSLAPAVFAASSNSFPVSVIQVNSAGQREYLLCPHEFGVFDVSYGRRS 1844  
Qy 1795 RTDDLKWSRLPLAPAYREPVLVTHFNSLEVIQIARSSAGTAPARAYLDTNPNRYLGPAL 1854  
Db 1845 RTDDLKWSRLPLAPAYREPVLVTHFNSLEVIQIARSSAGTAPARAYLDTNPNRYLGPAL 1904  
Qy 1855 SSGAIYLAASSYQDKRLVICCKGNLVKESGTEHHRGEPSTSSPNKRGPPTYNHEHITKRV 1914  
Db 1905 SSGAIYLAASSYQDKRLVICCKGNLVKESGTEHHRGEPSTSSPNKRGPPTYNHEHITKRV 1964  
Qy 1915 SSPAPPSPGSHPREPSTPHRY--REGTELRDOKSPGRPLEREKSPGRMLSTRERSPGR 1972  
Db 1965 SSPAPPSPGSHPREPSTPHRYRREGTELRDOKSPGRPLEREKSPGRMLSTRERSPGR 2024  
Qy 1973 LFDSSSRGRLPAGAVRTPLSQVNVKRGQSA 2002  
Db 2025 LFDSSSRGRLPAGAVRTPLSQVNVKQVMDQSS 2054

RESULT 12

ABB81928  
ID ABB81928 standard; protein; 1958 AA.

XX

AC ABB81928;

XX

DT 10-OCT-2002 (first entry)  
XX Human kinase #2.  
XX Human; kinase; enzyme; serine-threonine kinase; nootropic; cytostatic;  
KW Citron rho-interacting kinase; gene therapy; mental disorder; cancer.  
XX Homo sapiens.  
XX WO200259325-A2.  
XX 01-AUG-2002.  
XX 20-DEC-2001; 2001WO-US050497.  
XX 27-DEC-2000; 2000US-02598335P.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Yu X, Miranda M, Friddle CJ;  
XX WPI; 2002-599796/64.  
XX N-PSDB; ABQ78871.  
XX Novel polynucleotide encoding human proteins that are structurally  
PT similar to animal kinases, useful for drug screening, diagnosis, in gene  
PT therapy of disorders and diseases e.g. cancer and pharmacogenomic  
PT applications.  
XX Claim 4; Page 46-50; 50pp; English.  
XX The invention relates to a novel human protein that shares structural  
CC similarity with animal kinases, including serine-threonine kinases,  
CC particularly citron rho-interacting kinases. The proteins of the  
CC invention have nootropic and cytostatic activity. The polynucleotides may  
CC have a use in gene therapy. The encoded novel polypeptides are useful for  
CC generating antibodies, as reagents in diagnostic assays, for identifying  
CC other cellular gene products related to NHP and as reagents in assays for  
CC screening for compounds that are useful in the treatment of mental,  
CC biological or medical disorders and diseases including cancer. The  
CC sequence represents a novel human kinase of the invention  
XX Sequence 1958 AA;  
Query Match 90.4%; Score 9487.5; DB 5; Length 1958;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 1877; Conservative 1; Mismatches 1; Indels 81; Gaps 3;  
QY 1 MLKFKYGARNPLDAGAAEPIASRRLNLFQGGKPPMTQQQMSPLSREGILDALFVLPE 60  
DB 1 MLKFKYGARNPLDAGAAEPIASRRLNLFQGGKPPMTQQQMSPLSREGILDALFVLPE 60  
QY 61 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGGCHFAEVQVVRKATG 120  
DB 61 ECSQPALMKIKHVSFVRKYSDTIAELQELQPSAKDFEVRSLVGGCHFAEVQVVRKATG 120  
QY 121 DIYAMKMKKALLAQEQVSFEERNILSRSTSPWIPOLQYAFQDKNHLVLMVEYQPGG 180  
DB 121 DIYAMKMKKALLAQEQVSFEERNILSRSTSPWIPOLQYAFQDKNHLVLMVEYQPGG 180  
QY 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIKPNILVDRTGHLKLVDF 240  
DB 181 DLLSLNRYEDQDENLQFYLAELILAVSHVLMGYVHRDIKPNILVDRTGHLKLVDF 240  
QY 241 GSAAKNSNKNVNAKLPIGTPTDYMAPEVLTVNMGDKGTGYGLDCCDWSVGVIAYEMIYGR 300  
DB 241 GSAAKNSNKNVNAKLPIGTPTDYMAPEVLTVNMGDKGTGYGLDCCDWSVGVIAYEMIYGR 300  
QY 301 SPFAECTSARTNNIMNFRLKFPDDPKVSDFLDLTQSLCCKGKERLKPEGLCCHPFF 360  
DB 301 SPFAECTSARTNNIMNFRLKFPDDPKVSDFLDLTQSLCCKGKERLKPEGLCCHPFF 360  
QY 361 SKIDWNNIRNSPPFPVTLKSDDDTSNFDPEPKNSWSSSPCQLSPGFSGEELPFVGFs 420

DB 361 SKIDWNNIRNSPPFPVTLKSDDDTSNFDPEPKNSWSSSPCQLSPGFSGEELPFVGFs 420  
QY 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKQMEQMTLHRVRS 480  
DB 421 YSKALGILGRSESVVSGLDSPAKTSSMEKKLLIKSKELQDSQDKCHKQMEQMTLHRVRS 480  
QY 481 EVEAVLSQKEVELKASRTORSLLSLEODLATVITECSSLKRSLQEARMEVSEDDKALQLH 540  
DB 481 EVEAVLSQKEVELKASRTORSLLSLEODLATVITECSSLKRSLQEARMEVSEDDKALQLH 540  
QY 541 DIREQSRKLOEIKBOEYQAOVEENRMLMMNQLLEEDLVSARRRSDLYESELRESRLAAEFK 600  
DB 541 DIREQSRKLOEIKBOEYQAOVEENRMLMMNQLLEEDLVSARRRSDLYESELRESRLAAEFK 600  
QY 601 RKATECOHKLLKADQCKPVGVEYAKLEKINAGQKIQLQELQKLEKAVKASTEATELLQ 660  
DB 601 RKATECOHKLLKADQCKPVGVEYAKLEKINAGQKIQLQELQKLEKAVKASTEATELLQ 660  
QY 649 ----AKERAARELEFKLQNRDSDSEGIKKLVEABERHSLNKKVRLTMMERENRLKDD 704  
DB 661 NIRQAKERAARELEFKLQNRDSDSEGIKKLVEABERHSLNKKVRLTMMERENRLKDD 720  
QY 705 IOTXSQOIQQVADKILELEBKHREAQVSAQHLEVHLKQKEQHYBEKIKVLDNQIKKDLAD 764  
DB 721 IOTXSQOIQQVADKILELEBKHREAQVSAQHLEVHLKQKEQHYBEKIKVLDNQIKKDLAD 780  
QY 765 KETLENMQRHEEAEHEKGIKILSEQKAMINAMDSKIRSLQRIIVELSEANKLAANSLSFT 824  
DB 781 KETLENMQRHEEAEHEKGIKILSEQKAMINAMDSKIRSLQRIIVELSEANKLAANSLSFT 840  
QY 825 QRNKAQEMISELRQOKFYLETOAGKLEAQNRLKEEQLKISHQDSDKNRLLLELTRL 884  
DB 841 QRNKAQEMISELRQOKFYLETOAGKLEAQNRLKEEQLKISHQDSDKNRLLLELTRL 900  
QY 885 REVSLHEEOKLEKRLTLOLSLOERESQALQARAALLESOLFQAKTELEETTAAE 944  
DB 901 REVSLHEEOKLEKRLTLOLSLOERESQALQARAALLESOLFQAKTELEETTAAE 960  
QY 945 EEEIQAULTAHRDEIQRFKDALRNSCTVITDLEEQNLQTEDNAPLNQNFYLSKQLEDEAS 1004  
DB 961 EEEIQAULTAHRDEIQRFKDALRNSCTVITDLEEQNLQTEDNAPLNQNFYLSKQLEDEAS 1020  
QY 1005 GANDEIVQLRSEVDHLRREITEREMQLTQKOTMEALKTCTMLEEQVMDLEALNDELLE 1064  
DB 1021 GANDEIVQLRSEVDHLRREITEREMQLTQKOTMEALKTCTMLEEQVMDLEALNDELLE 1080  
QY 1065 KERQWEAWRSVLGDEKSOFCRVRELQRMOLDTEKQSRARADQRTESRQVVELAVKEHKA 1124  
DB 1081 KERQWEAWRSVLGDEKSOFCRVRELQRMOLDTEKQSRARADQRTESRQVVELAVKEHKA 1140  
QY 1125 EILALQALKEQKLKABSLSDKLNLEKHAMLEMNARSLOQKLETERELKQRLLEEQAK 1184  
DB 1141 EILALQALKEQKLKABSLSDKLNLEKHAMLEMNARSLOQKLETERELKQRLLEEQAK 1200  
QY 1185 LQQQMDLQKNHIFRLTOGLOEALDRADLLKTERSDLEYOLENIQVLYSHEKVKMEGTISQ 1244  
DB 1201 LQQQMDLQKNHIFRLTOGLOEALDRADLLKTERSDLEYOLENIQVLYSHEKVKMEGTISQ 1260  
QY 1245 QTKLIDFLQAKMPAKKKGLFSRRKEDPALPTQVPLQVYNELKLALKEKAKCAELEEA 1304  
DB 1261 QTKLIDFLQAKMPAKKKK-----VPLQVYNELKLALKEKAKCAELEEA 1305  
QY 1305 LQKTRIELRSARBEAAHRKATDHPSTPATACQITAMSAIVSPHOPSAMSLAPPSS 1364  
DB 1306 LQKTRIELRSARBEAAHRKATDHPSTPATACQITAMSAIVSPHOPSAMSLAPPSS 1365  
QY 1365 RRKESSTPEFSRRLKERMHNIPIHRFNVGLNNRATKCAVCLDTVHFGROASKLECOVM 1424  
DB 1366 RRKESSTPEFSRRLKERMHNIPIHRFNVGLNNRATKCAVCLDTVHFGROASKLECOVM 1425  
QY 1425 CHPKCSTCLPATGCLPABYATHTEAFCDKXNSPGLQTHKEPSSSLHLEGMKVPNNKR 1484

Db 1426 CHPKCTCLPATCGLPAEYATHFTEAFORDKMNKSPGLQTKPESSSLHLEGWVKVPRNNKR 1485  
QY 1485 GQGQWRKYIVLGGSKVLIYDNEARAGORPVEEFELCLPDGQVSIHGAVGASELANTAK 1544  
Db 1486 GQGQWRKYIVLGGSKVLIYDNEARAGORPVEEFELCLPDGQVSIHGAVGASELANTAK 1545  
QY 1545 A-----BKAADAKL 1554  
Db 1546 ADVPYILKMBESHPTTCWPGRTLYLLAPSPDKQRWVTALLESVYAGGRVSRKAEADAKL 1605  
QY 1555 LGSNLLKLEGDDRLDNWCTLPSFDQVVLVGTBEGYALVKNLSLTHVPGIGAVFOIYII 1614  
Db 1606 LGSNLLKLEGDDRLDNWCTLPSFDQVVLVGTBEGYALVKNLSLTHVPGIGAVFOIYII 1665  
QY 1615 KDLEKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIFFAVKGCFLFGAGKIEN 1674  
Db 1666 KDLEKLLMIAGEERALCLVDVKVKQSLAQSHLPAQPDISPNIFFAVKGCFLFGAGKIEN 1725  
QY 1675 GLCICAMPSKVILRYNENLSKYCKIELETSSEPCSCIFHTNYSILIGTNKPYEIDMKQ 1734  
Db 1726 GLCICAMPSKVILRYNENLSKYCKIELETSSEPCSCIFHTNYSILIGTNKPYEIDMKQ 1785  
QY 1735 YTLLEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREXYLLCFHEFGVFDVSYGRRS 1794  
Db 1786 YTLLEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREXYLLCFHEFGVFDVSYGRRS 1845  
QY 1795 RTDDLKWSRLPAPAFREPFLVFTHNSLEVIETIQARSAGTAPARAYLDIPNRYLGPAL 1854  
Db 1846 RTDDLKWSRLPAPAFREPFLVFTHNSLEVIETIQARSAGTAPARAYLDIPNRYLGPAL 1905  
QY 1855 SSGAIVLASSYQDKLRVICCKGNLVKESGTEHHRGPGTGR 1894  
Db 1906 SSGAIVLASSYQDKLRVICCKGNLVKESGTEHHRGPGTGR 1945

RESULT 13  
ADD48584  
ID ADD48584 standard; protein; 1619 AA.  
AC ADD48584;  
XX  
DT 29-JAN-2004 (first entry)  
DE Rat Protein AF039218, SEQ ID NO 14290.  
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX Rattus norvegicus.  
OS  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GENO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX GENBANK; AF039218.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.  
XX

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
CC injury (CCI) and spared nerve injury (SNI)).  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1619 AA;  
Query Match 71.7%; Score 7521; DB 7; Length 1619;  
Best Local Similarity 92.7%; Pred. No. 0;  
Matches 1500; Conservative 18; Mismatches 16; Indels 84; Gaps 5;  
QY 468 MEQMTLHRRVSEVAVLSQKEVELKASRTORSLLEODLATYITCSSLKRSLEQARME 527  
Db 1 MEQEMARLHRRVSEVAVLSQKEVELKASRTORSLLEODLATYITCSSLKRSLEQARME 60  
QY 528 VSQEDDKALQLLHDIRQSRKLOEIKQEYQAOVEEMRLMMNQLEEDLVSAARRSDIYES 587  
Db 61 VSQEDDKALQLLHDIRQSRKLOEIKQEYQAOVEEMRLMMNQLEEDLVSAARRSDIYES 120  
QY 588 ELRESRLAAEFKFKATECOHKLKAKDQKPEVGEYAKLEKINAEQOLKIQLQEKLEK 647  
Db 121 ELRESRLAAEFKFKATECOHKLKAKDQKPEVGEYAKLEKINAEQOLKIQLQEKLEK 180  
QY 648 A-----AKERAERELEKLNQNRDSSSGIRKRLVEABERHSHLENKVKRL 691  
Db 181 AVKASTATALLQNIROAKERARELEKLNHRDSSSGIRKRLVEABERHSHLENKVKRL 240  
QY 692 ETMERRENRLKDDIQTQSQQIQMADKILELEKHREDAQVSAQHLEVLHKKQCHYEKI 751  
Db 241 ETMERRENRLKDDIQTQSQQIQMADKILELEKHREDAQVSAQHLEVLHKKQCHYEKI 300  
QY 752 KVLNDOIKKDLADKETLENMQRHEEAHEKGLTSPKAMINAMDSKIRSLQRIVELS 811  
Db 301 KVLNDOIKKDLADKETLENMQRHEEAHEKGLTSPKAMINAMDSKIRSLQRIVELS 360  
QY 812 EANKLAANSILFTORNKKAQEMISELRQOKFYLETQAGKLEAQNRLKLEPQKISHQDH 871  
Db 361 EANKLAANSILFTORNKKAQEMISELRQOKFYLETQAGKLEAQNRLKLEPQKISHQDH 420  
QY 872 SDKRRLLETRLEVSLEHEEQKLELKRQITLQLSLOERESQLTALQARAALSQLR 931  
Db 421 SDKRRLLETRLEVSLEHEEQKLELKRQITLQLSLOERESQLTALQARAALSQLR 480  
QY 932 QAKTELEBTTAAEBEEIQAALTAHRDEIQRFKFDALRNSCTVITDLEEQNLQTEDNAELNN 991  
Db 481 QAKTELEBTTAAEBEEIQAALTAHRDEIQRFKFDALRNSCTVITDLEEQNLQTEDNAELNN 540  
QY 992 QNFYLSKQDLASGANDIEIVQLRSEVDHLRREITEREMQLTQSKOTWEALKTTCTMLEEQ 1051  
Db 541 QNFYLSKQDLASGANDIEIVQLRSEVDHLRREITEREMQLTQSKOTWEALKTTCTMLEEQ 600

Qy	1052	VMDLEALNDELLEKXEROWEAWRSVLGDEKQOFCRVRBELQSMILDTKQSRARADQRITES	1111
Db	601	VMDLEALNDELLEKXEROWEAWRSVLGDEKQOFCRVRBELQSMILDTKQSRARADQRITES	660
Qy	1112	RQVVELAVKEHKABILAQALKEQKLBKSLSDKLNDLEKKGHAMLEMNARSLOQKLETE	1171
Db	661	RQVVELAVKEHKABILAQALKEQKLBKSLSDKLNDLEKKGHAMLEMNARSLOQKLETE	720
Qy	1172	RELKORLELEQAKLOQOMDLQKNHIFRLTQGLQALRADLLKTERSDLEQLENIOVLY	1231
Db	721	RELKORLELEQAKLOQOMDLQKNHIFRLTQGLQALRADLLKTERSDLEQLENIOVLY	780
Qy	1232	SHEKVMKEGTISQOTKLIDFLQAKMDQPAKKKGILFSRRKEDPALPTQVLPQYNELKIAL	1291
Db	781	SHEKVMKEGTISQOTKLIDFLQAKMDQPAKKKGILFSRRKEDPALPTQVLPQYNELKIAL	840
Qy	1292	EKEKARCAEUEEALQKTRIEILRSAREEAAHRKATDHPHPTATPARQOIAWSAIVRSPH	1351
Db	841	EKEKARCAEUEEALQKTRIEILRSAREEAAHRKATDHPHPTATPARQOIAWSAIVRSPH	900
Qy	1352	QPSAMSLAPPPSSRRKSSSTPEEPSRRLKERMHHNIPHRFNVLGNMRAATCAVCLDTVHF	1411
Db	901	QPSAMSLAPPPSSRRKSSSTPEEPSRRLKERMHHNIPHRFNVLGNMRAATCAVCLDTVHF	960
Qy	1412	GRQASKCLEQVMCHPKCSTCLPATCGLPAEYATHFTEAFCRDKMNSPGQTKERPSSSLH	1471
Db	961	GRQASKCLEQVMCHPKCSTCLPATCGLPAEYATHFTEAFCRDKVSPGQSKERPSSSLH	1020
Qy	1472	LEGMMKVPNRNKGQQGWDRKYIVLEGSKVLIYDNEAREAGQRPVEBEFCLPDGDVSIH	1531
Db	1021	LEGMMKVPNRNKGQQGWDRKYIVLEGSKVLIYDNEAREAGQRPVEBEFCLPDGDVSIH	1080
Qy	1532	GAVGASELANTAKA-----	1545
Db	1081	GAVGASELANTAKADVPIYKWMESHPTTQWPGRTLYLLAPSPDKQWTALESVVAGG	1140
Qy	1546	-----EKAEADA-----	1586
Db	1141	RVBREKAEADAAMDCTSERLPVWVEKLLGNSLLKEGDELDMNCTLPFSQOVLVVGTE	1200
Qy	1587	EGYALNVLKNSLTHVFGI GAVFOIYIIXDLEKLLMAGEERALCLVDVKVKVQSQAQSH	1646
Db	1201	EGYALNVLKNSLTHI FGI GAVFOIYIIXDLEKLLMAGEERALCLVDVKVKVQSQAQSH	1260
Qy	1647	LPQAQPDISPNI FEAVKQCHLFGACKIENGICICAMPSKVVIIRYNNENLSKYCIRKEIET	1706
Db	1261	LPQAQPDISPNI FEAVKQCHLFGACKIENSJCICAMPSKVVIIRYNNENLSKFCIRKEIET	1320
Qy	1707	SEPCSCIHFTNYSILITGNKFYEIDMKQYTLBEEFLDKNDHSLAPAVFAAASNSFPVSIQ	1766
Db	1321	SEPCSCIHFTNYSILITGNKFYEIDMKQYTLBEEFLDKNDHSLAPAVFASSTNSFPVSIQ	1380
Qy	1767	VNSAGQREBYLLCFHERGVFDVSGRRSRTDDLKWSRLPLAFAYREBYLIVTHFNSLEVI	1826
Db	1381	ANSFGQREBYLLCFHERGVFDVSGRRSRTDDLKWSRLPLAFAYREBYLIVTHFNSLEVI	1440
Qy	1827	EIQARSSAGTPARAYLIDIPNRYLGPAPISGAIYASSYQDKLRVICCKGNLVKESGTEH	1886
Db	1441	EIQARSSILGTPARAYLIPNRYLGPAPISGAIYASSYQDKLRVICCKGNLVKESGTEQ	1500
Qy	1887	HRCGPSTRSPNKGPBTYNEHITKRVASSAPPEGSHPRESTPHRY--REGRIELRR	1944
Db	1501	HRVPSTGR--GPNKKGPTTYNEHITKRVASSAPPEGSHPRESTPHRYDRDGRTELR	1559
Qy	1945	DKSGRPLEREKPSGRMLSTRRESRQRLFPEDSSRGRLPAGAVRTPUSQVNMKRGQSA	2002
Db	1560	DKSGRPLEREKPSGRMLSTRRESRQRLFPEDSSRGRLPAGAVRTPUSQVNMKWDQSS	1617

ADD46516;  
29-JAN-2004 (first entry)  
Rat Protein AF039218, SEQ ID NO 12299.  
Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
Rattus norvegicus.  
WO2003016475-A2.  
27-FEB-2003.  
14-AUG-2002; 2002WO-US025765.  
14-AUG-2001; 2001US-0312147P.  
01-NOV-2001; 2001US-0346382P.  
26-NOV-2001; 2001US-0333347P.  
(GEHO ) GEN HOSPITAL CORP.  
(FARB ) BAYER AG.  
Woolf C, D'urso D, Befort K, Costigan M;  
WPI; 2003-268312/26.  
GENBANK; AF039218.  
New composition comprising two or more isolated polypeptides, useful for  
preparing a medicament for treating pain in an animal.  
Claim 1; Page; 1017pp; English.  
The invention discloses a composition comprising two or more isolated rat  
or human polynucleotides or a polynucleotide which represents a fragment,  
derivative or allelic variation of the nucleic acid sequence. Also  
claimed are a vector comprising the novel polynucleotide, a host cell  
comprising the vector, a method for identifying a nucleotide sequence  
which is differentially regulated in an animal subjected to pain and a  
kit to perform the method, an array, a method for identifying an agent  
that increases or decreases the expression of the polynucleotide sequence  
that is differentially expressed in neuronal tissue of a first animal  
subjected to pain, a method for identifying a compound which regulates  
the expression of a polynucleotide sequence which is differentially  
expressed in an animal subjected to pain, a method for identifying a  
compound that regulates the activity of one or more of the  
polynucleotides, a method for producing a pharmaceutical composition, a  
method for identifying a compound or small molecule that regulates the  
activity in an animal of one or more of the polypeptides given in the  
specification, a method for identifying a compound useful in treating  
pain and a pharmaceutical composition comprising the one or more  
polypeptides or their antibodies. the polynucleotide or the compound that  
modulates its activity is useful for preparing a medicament for treating  
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
therapy). The sequence presented is a rat protein (shown in Table 2 of  
the specification) which is differentially expressed during pain. Note:  
The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic form directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences.  
Sequence 1619 AA;  
SQ

RESULT 14  
ADD46616  
ID ADD4

QY 528 VSQEDKALQLLHDIREQSRKQEIKEQYQAOVEEMRLMNQLEEDLVASRRSDLYES 587  
 DB 61 VSQEDKALQLLHDIREQSRKQEIKEQYQAOVEEMRLMNQLEEDLVASRRSDLYES 120  
 QY 588 ELRESLAAEFKPKATECHQKLLKAKDOGKPEVGEYAKLEKINABOOLKIOLEKLEK 647  
 DB 121 ELRESLAAEFKPKANECHQKLLKAKDOGKPEVGEYAKLEKINABOOLKIOLEKLEK 180  
 QY 648 A-----AKERARELEKLEQNREDSSEGIRKKLVABEERRHLENKVKRL 691  
 DB 181 AVFASTEATELLQNIQAKERARELEKLEHNREDSSEGIRKKLVABEERRHLENKVKRL 240  
 QY 692 ETWERENRKDDIQTKSQIQOQADKILEBEKHREAQVSACHLEVLHUKQEQHYEKEI 751  
 DB 241 ETWERENRKDDIQTKSQIQOQADKILEBEKHREAQVSACHLEVLHUKQEQHYEKEI 300  
 QY 752 KVLNDQIKDLADKETLENMORHEEAEHEKGKILSEOKAMNAMDSKIRSLRQIVELS 811  
 DB 301 KVLNDQIKDLADKESLETMQHHEEAEHEKGKILSEOKAMNAMDSKIRSLRQIVELS 360  
 QY 812 EANKLAANSLFTQNNKAQEMISELROQKPYLETQAGKLEAQRKLEBEQLEKISHQDH 871  
 DB 361 EANKLAANSLFTQNNKAQEMISELROQKPYLETQAGKLEAQRKLEBEQLEKISHQDH 420  
 QY 872 SDKNRLELETRREYLSHEEOKLEKQLELOLSLOERSQLTALQAAALLESQJR 931  
 DB 421 SDKNRLELETRREYLSHEEOKLEKQLELOLSLOERSQLTALQAAALLESQJR 480  
 QY 932 QAKTELEETAAEABEIEIQAHTARDEIQKFDALRNSCTVITDLEBQLNQLTDEDAELNN 991  
 DB 481 QAKTELEETAAEABEIEIQAHTARDEIQKFDALRNSCTVITDLEBQLNQLTDEDAELNN 540  
 QY 992 QNPLYSKQLDEASGANDIEVQLRSEVDHLRREITEREMQLTTSQKQTMELKTTCTMLBEQ 1051  
 DB 541 QNPLYSKQLDEASGANDIEVQLRSEVDHLRREITEREMQLTTSQKQTMELKTTCTMLBEQ 600  
 QY 1052 VMDLEALNDELLEKQWEAWRSVLGDEKSQCEVRVRELQRMLDTEKQSRARADQRTES 1111  
 DB 601 VMDLEALNDELLEKQWEAWRSVLGDEKSQCEVRVRELQRMLDTEKQSRARADQRTES 660  
 QY 1112 RQVVELAVHEKHAIEIALQALKEQKKAESLSDKLNLEKHAMLEMMARSLOQKLETE 1171  
 DB 661 RQVVELAVHEKHAIEIALQALKEQKKAESLSDKLNLEKHAMLEMMARSLOQKLETE 720  
 QY 1172 RELKORLLEBQAKLOQMDLQKNHIFRLTQGLQEALDRADLLKTERSDEYOLENIQVLY 1231  
 DB 721 RELKORLLEBQAKLOQMDLQKNHIFRLTQGLQEALDRADLLKTERSDEYOLENIQVLY 780  
 QY 1232 SHEKVRVEGTISQOTKLIDFLQAKMDQPAKKKGLFSRKEKDPALPTQVPLQYNELKIAL 1291  
 DB 781 SHEKVRVEGTISQOTKLIDFLQAKMDQPAKKKGLFSRKEKDPALPTQVPLQYNELKIAL 840  
 QY 1292 EKEKARCAELEALOKTRLELSAREEAAHRKATDPHSTPATARQQIAMSALVRSPEH 1351  
 DB 841 EKEKARCAELEALOKTRLELSAREEAAHRKATDPHSTPATARQQIAMSALVRSPEH 900  
 QY 1352 QPSAMSLAPPSSRRKESSTPEEFSRRLKERMHNNIPHRFNVGLNMRATKCAVCLDTVHF 1411  
 DB 901 QPSAMSLAPPSSRRKESSTPEEFSRRLKERMHNNIPHRFNVGLNMRATKCAVCLDTVHF 960  
 QY 1412 GROASKLEQVWCHPKCSTCLPATCGLPAEYATHEFAFCDKXNSPGLQTKESPSSLIH 1471  
 DB 961 GROASKLEQVWCHPKCSTCLPATCGLPAEYATHEFAFCDKXNSPGLQTKESPSSLIH 1020  
 QY 1472 LEGWMKVPRNNKGQGWDRKYIVLFGSKVLYIDNEAREAGQRPVEFELCLPDGDVSIH 1531  
 DB 1021 LEGWMKVPRNNKGQGWDRKYIVLFGSKVLYIDNEAREAGQRPVEFELCLPDGDVSIH 1080  
 QY 1532 GAVGASELANTAKA----- 1545  
 DB 1081 GAVGASELANTAKADVPYILKXESHPTTCWFQRTLYLLAPSPFDKQRWVTALESVWAGG 1140  
 QY 1546 -----EKAEDA-----KLLGNSLLKLEGGDRDLMNCTLPFSQVVLVGE 1586

DB 1141 RVSREKAEADAAMDCTCERLPMVVEKLLGNSLLKLEGGDRDLMNCTLPFSQVVLVGE 1200  
 QY 1587 EGLYALNVLKNSLTHVPVIGAVFOIYIIKOLEKLLMIAGEERALCLVDVKKVQSLAQSH 1646  
 DB 1201 EGLYALNVLKNSLTHVPVIGAVFOIYIIKOLEKLLMIAGEERALCLVDVKKVQSLAQSH 1260  
 QY 1647 LPAQPDISPNIPEAVKCHLFGAGKLENGLCICAAMPSKVILRYNENLSKYCIKRIEIT 1706  
 DB 1261 LPAQPDISPNIPEAVKCHLFGAGKLENGLCICAAMPSKVILRYNENLSKYCIKRIEIT 1320  
 QY 1707 SEPSCCHTFNYSYLLIGTNKFYEIDMKQYLTLEEFLLKNDHSLAPAFVAAASNSFPVSIQ 1766  
 DB 1321 SEPSCCHTFNYSYLLIGTNKFYEIDMKQYLTLEEFLLKNDHSLAPAFVAAASNSFPVSIQ 1380  
 QY 1767 VNSAQREEVYLLCFHEFGVFDVSYGRSRTDDDKWSRLPLAFAYREPYLVTFHNSLEVI 1826  
 DB 1381 ANSTGQREEVYLLCFHEFGVFDVSYGRSRTDDDKWSRLPLAFAYREPYLVTFHNSLEVI 1440  
 QY 1827 EIQRSSAGTTPARAYLDIPNPRYLGPALSSGAYLASSYQDKLRVICCKGNLVKESGTEH 1886  
 DB 1441 EIQRSSAGTTPARAYLDIPNPRYLGPALSSGAYLASSYQDKLRVICCKGNLVKESGTEH 1500  
 QY 1887 HRGFTSRSSPNKRGPPPTNEHITKRVASSPAPPEGSHPREPSTPHRY--REGRTELRR 1944  
 DB 1501 HRVFTSR--SPNKGPPPTNEHITKRVASSPAPPEGSHPREPSTPHRYDRDREGRTELRR 1559  
 QY 1945 DXSGRPLEKSPGRMLSTRERSPGRLFEDESRGLPAGAVRTPLSQVKNKGQOSA 2002  
 DB 1560 DXSGRPLEKSPGRMLSTRERSPGRLFEDESRGLPAGAVRTPLSQVKNKGQOSA 1617

RESULT 15  
 AAB43359 ID AAB43359 standard; protein; 1286 AA.  
 XX AAB43359;  
 AC AAB43359;  
 DT 08-FEB-2001 (first entry)  
 XX Human ORFX ORF3123 polypeptide sequence SEQ ID NO:6246.  
 DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malacia; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX Homo sapiens.  
 XX WO200058473-A2.  
 PN 05-OCT-2000.  
 PD 31-MAR-2000; 2000WO-US008621.  
 XX 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX (CURA-) CURAGEN CORP.  
 PA Shimkets RA, Leach M;  
 PI

XX WPI; 2000-602362/57.  
DR N-PSDB; AAC77568.  
XX  
PT Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease.  
XX  
PS Claim 11; Page 5433-5436; 5507pp; English.  
XX  
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
CC antiparasitic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
CC anticovulant; antiarthritic; immunosuppressant; immunostimulant;  
CC cardiant; thrombolytic; coagulant; vasotrophic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
SQ Sequence 1286 AA;  
Query Match 59.3%; Score 6223.5; DB 3; Length 1286;  
Best Local Similarity 94.7%; Pred. No. 0;  
Matches 1231; Conservative 1; Mismatches 3; Indels 65; Gaps 2;  
QY 753 VLDNOIKDLADKTELENNMQHHEAEKGIKILSEQKAMINAMDSKIRSEQRIVELSE 812  
DB 1 VLDNOIKDLADKTELENNMQHHEAEKGIKILSEQKAMINAMDSKIRSEQRIVELSE 60  
QY 813 ANKLAANSLFTQRMKAQEEIMSELROQKFYLETQAGKLEAQNKLEEQLEKISHODHS 872  
DB 61 ANKLAANSLFTQRMKAQEEIMSELROQKFYLETQAGKLEAQNKLEEQLEKISHODHS 120  
QY 873 DKNRLELETRLVSLHEEOKLEKQLTELQLSOERESQTLQAAAALLESQLRQ 932  
DB 121 DKNRLELETRLVSLHEEOKLEKQLTELQLSOERESQTLQAAAALLESQLRQ 180  
QY 933 AKTELEETTABAEIEIQALTARHDEIQKFDALRNSCTVITDLEEQNLQTDNAELNNQ 992  
DB 181 AKTELEETTABAEIEIQALTARHDEIQKFDALRNSCTVITDLEEQNLQTDNAELNNQ 240  
QY 993 NPYLSQKLDASGANDIEIVQLRSEVDHLRREITEREMOLTSQKQTMELAKTCTMLEEQV 1052  
DB 241 NPYLSQKLDASGANDIEIVQLRSEVDHLRREITEREMOLTSQKQTMELAKTCTMLEEQV 300  
QY 1053 MDLEALNDELLEKQEWAEWSVLGDEKSOECRVRELQRMIDTEKQSRADQRIETSR 1112  
DB 301 MDLEALNDELLEKQEWAEWSVLGDEKSOECRVRELQRMIDTEKQSRADQRIETSR 360  
QY 1113 QVVELAVKHEKAEIILALQALKEQKKAESLSDKLNLEKKHAMLENNARSLOQKLETER 1172  
DB 361 QVVELAVKHEKAEIILALQALKEQKKAESLSDKLNLEKKHAMLENNARSLOQKLETER 420  
QY 1173 ELKQRLLEBEQAKLQOQMDLQKNHIFRLTQGLQOALDRADLLKTERSULEYOLENIQVLYS 1232  
DB 421 ELKQRLLEBEQAKLQOQMDLQKNHIFRLTQGLQOALDRADLLKTERSULEYOLENIQVLYS 480  
QY 1233 HEKVKWEGTISQOTKLIDFLQAKMDQPAKKKGLFRRKEDPALPTQVPLQYNELKLALE 1292  
DB 481 HEKVKWEGTISQOTKLIDFLQAKMDQPAKKK-----VPLQYNELKLALE 525

QY 1293 KEKARCAELBEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQOIAISAIVRSEHQ 1352  
DB 526 KEKARCAELBEALQKTRIELRSAREEAAHRKATDHPHPSTPATARQOIAISAIVRSEHQ 585  
QY 1353 PSAMSLIAPSSRRKESSTPEEFSRRLKERMHNHIFHRFVGLNMRATKCAVCLDTVHFG 1412  
DB 586 PSAMSLIAPSSRRKESSTPEEFSRRLKERMHNHIFHRFVGLNMRATKCAVCLDTVHFG 645  
QY 1413 RQASKCLECOVMCHPKCSTCLPATCGIPAEVATHTFAFCRDKNVSPGLQTKPSSSLHL 1472  
DB 646 RQASKCLECOVMCHPKCSTCLPATCGIPAEVATHTFAFCRDKNVSPGLQTKPSSSLHL 705  
QY 1473 EGMKVPNNKRGQGWDRKYIIVLEGSKVLIYDNEAREAGQRPVEFELCLPDGDSIHG 1532  
DB 706 EGMKVPNNKRGQGWDRKYIIVLEGSKVLIYDNEAREAGQRPVEFELCLPDGDSIHG 765  
QY 1533 AVGASELANATA----- 1545  
DB 766 AVGASELANATAKADVPVILKWEHSHHTCPWPGRTLYLLAESPFPKQRWVTALESVAVGR 825  
QY 1546 ---EKAEADAKLLGNSLLKLEGGDRDLDNMCTLPFSQVVLVGTGEGLYALNVLKNSLTHV 1602  
DB 826 VSREKAEADAKLLGNSLLKLEGGDRDLDNMCTLPFSQVVLVGTGEGLYALNVLKNSLTHV 885  
QY 1603 PGIGAVPQIYIIXDLEKLMIAEERALCLVDVKKVQSLAQSHLPAQDIPSNIFEAVK 1662  
DB 886 PGIGAVPQIYIIXDLEKLMIAEERALCLVDVKKVQSLAQSHLPAQDIPSNIFEAVK 945  
QY 1663 GCHLFGAGKIEGLICTCAAMPKVVILRYNENLSKYCIKEIETSPSCSIHFTNYSILI 1722  
DB 946 GCHLFGAGKIEGLICTCAAMPKVVILRYNENLSKYCIKEIETSPSCSIHFTNYSILI 1005  
QY 1723 GTNKFYEIDMKQYTLBEFLDKNDHSLAPAVFAASSNSFPVSIYQVNSAGQREYLLCFHE 1782  
DB 1006 GTNKFYEIDMKQYTLBEFLDKNDHSLAPAVFAASSNSFPVSIYQVNSAGQREYLLCFHE 1065  
QY 1783 FGVFVDSYGRSRTDCLKWSRLPLAFAYREPYLFTVTHNSLEVIETQABSAGTAPAYL 1842  
DB 1066 FGVFVDSYGRSRTDCLKWSRLPLAFAYREPYLFTVTHNSLEVIETQABSAGTAPAYL 1125  
QY 1843 DIPNRYLGPASSGAIYLAASSYQDLRVTCCKGNLVKESGTEHHRGPFSTRSSPNKRG 1902  
DB 1126 DIPNRYLGPASSGAIYLAASSYQDLRVTCCKGNLVKESGTEHHRGPFSTRSSPNKRG 1185  
QY 1903 PTYNEHITKRVASSPAPPQPSHPREPSTPHRYREGRTELRRDKSPGRPLEREKSPGRML 1962  
DB 1186 PTYNEHITKRVASSPAPPQPSHPREPSTPHRYREGRTELRRDKSPGRPLEREKSPGRML 1245  
QY 1963 STRRERSPGRLFEDSSGRPLPAGAVRTPLSQVKNKGQSA 2002  
DB 1246 STRRERSPGRLFEDSSGRPLPAGAVRTPLSQVKNKGQSA 1285

Search completed: July 3, 2004, 10:14:52  
Job time : 111 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 18:47:57 ; Search time 24219 Seconds  
(without alignments)  
11765.024 Million cell updates/sec

Title: US-10-017-216-1  
Perfect score: 6574  
Sequence: 1 agagccgcagtgaggagat.....atcgagaatgtaggtttaga 6574

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pin.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_man.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sv.\*
- 39: em\_hgo\_hum.\*
- 40: em\_hgo\_mus.\*
- 41: em\_hgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	6574	100.0	6574	6	AX429512	AX429512 Sequence
2	6159	93.7	6159	6	AX429514	AX429514 Sequence
3	5684.4	86.5	6298	6	AX504254	AX504254 Sequence
4	5668	86.2	8603	6	AX671108	AX671108 Sequence
5	5666.6	86.2	6156	6	AX671112	AX671112 Sequence
6	5666.2	86.2	6165	6	AX671105	AX671105 Sequence
7	5661.4	86.1	6165	6	AX574425	AX574425 Sequence
8	5650.4	86.0	6159	6	AX166510	AX166510 Sequence
9	5631	85.7	6189	6	AX503780	AX503780 Sequence
10	5629.8	85.6	6201	6	AX503778	AX503778 Sequence
11	5552.8	84.5	8576	9	AY257469	AY257469 Homo sapi
12	5358	81.5	5877	6	AX574427	AX574427 Sequence
13	4632.4	70.5	6954	10	AF086824	AF086824 Mus muscu
14	3602.6	54.8	5952	10	AF039218	AF039218 Rattus no
15	3593.8	54.7	4967	10	AF070066	AF070066 Mus muscu
16	3546.4	53.9	5019	10	MM039904	U39904 Mus musculu
17	2436.8	37.1	5261	6	AX671110	AX671110 Sequence
18	2436.8	37.1	5261	9	AB023166	AB023166 Homo sapi
19	2187.4	33.3	3495	10	BC051165	BC051165 Mus muscu
20	2187.4	33.3	3506	10	BC023775	BC023775 Mus muscu
21	1406.4	21.4	2066	6	AX642956	AX642956 Sequence
22	1397.8	21.3	1515	6	AX253937	AX253937 Sequence
23	1395	21.2	1485	6	AX671037	AX671037 Sequence
24	1395	21.2	1765	6	AX671044	AX671044 Sequence
25	1087.8	16.5	2380	6	AX671042	AX671042 Sequence
26	1087.8	16.5	2380	10	AF086823	AF086823 Mus muscu
27	1083.4	16.5	1386	10	AF070065	AF070065 Rattus no
28	768	11.7	1133	6	AB411948	AB411948 Sequence
29	648	9.9	1058	6	AX053315	AX053315 Sequence
30	592	9.0	817	6	AX053416	AX053416 Sequence
31	570	8.7	2390	9	AF022116	AF022116 Homo sapi
32	570	8.7	2395	9	BC017671	BC017671 Homo sapi
33	570	8.7	2402	9	BC001007	BC001007 Homo sapi
34	570	8.7	2429	9	BC001056	BC001056 Homo sapi
35	570	8.7	2429	9	BC001823	BC001823 Homo sapi
36	570	8.7	2442	9	AK127820	AK127820 Homo sapi
37	570	8.7	136436	9	AC002563	AC002563 Human PAC
38	570	8.7	190514	2	AC026363	AC026363 Homo sapi
39	570	8.7	208686	9	AC079262	AC079262 Homo sapi
40	568.4	8.6	4122	9	HSM805803	BX537486 Homo sapi
41	544.8	8.3	2089	9	HSAMKPK6	U87276 Human AMP-a
42	544.8	8.3	2353	9	HSU3994	U83994 Human AMP-a
43	434	6.6	2284	9	AK123136	AK123136 Homo sapi
44	373.2	5.7	398	6	AX913973	AX913973 Sequence
45	373.2	5.7	398	6	BD049506	BD049506 Sequence

ALIGNMENTS

RESULT 1  
AX429512  
LOCUS AX429512 6574 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 1 from Patent WO0234896.  
ACCESSION AX429512  
VERSION AX429512.1 GI:21540791  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Kapeller-Libermann, R.  
TITLE 13245, a novel human myotonic dystrophy type protein kinase and  
uses therefor

QY	1	AGAGCGCGCAGTGGGGAGATGTTGAAGTTCAATATATGGACGCGGGAATCCTTTGGATGCT	60
Db	1	AGAGCGCGCAGTGGGGAGATGTTGAAGTTCAATATATGGACGCGGGAATCCTTTGGATGCT	60
QY	61	GGTGTGTGTGAACCATTTGCCAGACGGGCGCTCCAGGCTGAATCTGTCTTCCAGGGGAAA	120
Db	61	GGTGTGTGTGAACCATTTGCCAGCGGGCTCCAGGCTGAATCTGTCTTCCAGGGGAAA	120
QY	121	CCACCCCTTATGACTCAACACAGATGTCCTCTTTCCGAAAGGGATATTAGATGCC	180
Db	121	CCACCCCTTATGACTCAACACAGATGTCCTCTTTCCGAAAGGGATATTAGATGCC	180
QY	181	CTCTTTGTCTCTTTGAAGAAATGCAGTCAGCCTGCTCTGATGAAGATTAAACACCTGAGC	240
Db	181	CTCTTTGTCTCTTTGAAGAAATGCAGTCAGCCTGCTCTGATGAAGATTAAACACCTGAGC	240
QY	241	TACTTTGTCCGAGTATTTCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCA	300
Db	241	TACTTTGTCCGAGTATTTCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCA	300
QY	301	AAGGACTTCGAAGTCAGAAGTCTTTGAGTTGTGGTCACTTTCTGCTGAAGTCAGGTGGTA	360
Db	301	AAGGACTTCGAAGTCAGAAGTCTTTGAGTTGTGGTCACTTTCTGCTGAAGTCAGGTGGTA	360
QY	361	ACAGAGAAGCAACCGGGACATCTATGCTATGAAGTGAATGAAGAAGAGCTTTATTG	420
Db	361	ACAGAGAAGCAACCGGGACATCTATGCTATGAAGTGAATGAAGAAGAGCTTTATTG	420
QY	421	GCCGAGGACGAGTTTCATTTTGTAGGAAGACGGGAACATATTATCTCGAAGCAACAAC	480
Db	421	GCCGAGGACGAGTTTCATTTTGTAGGAAGACGGGAACATATTATCTCGAAGCAACAAC	480
QY	481	CGTGGATCCCCAATTACAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGATGAG	540
Db	481	CGTGGATCCCCAATTACAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGATGAG	540
QY	541	GAATATACGCTGTGAGGGGACCTGCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGAT	600
Db	541	GAATATACGCTGTGAGGGGACCTGCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGAT	600
QY	601	GAAGACCTGTACAGTTTTTACTAGCTAGCTGAATTTTGGCTCTTCACAGCGTTCACTGTG	660
Db	601	GAAGACCTGTACAGTTTTTACTAGCTAGCTGAATTTTGGCTCTTCACAGCGTTCACTGTG	660
QY	661	ATGGGATACGTGCAATCGAGACATCAAGCCCTGAGAAATTCCTGTTGACCGCACAGGAC	720
Db	661	ATGGGATACGTGCAATCGAGACATCAAGCCCTGAGAAATTCCTGTTGACCGCACAGGAC	720
QY	721	ATCAAGCTGGTGGATTTTGGATCTGCGCGGAAATGAATTCAAACAAGATGGTGAATGCC	780
Db	721	ATCAAGCTGGTGGATTTTGGATCTGCGCGGAAATGAATTCAAACAAGATGGTGAATGCC	780
QY	781	AAACTCTCCGATTTGGGACCCAGATTACATGGCTCTTGAAGTGTCTGATGAACCGG	840
Db	781	AAACTCTCCGATTTGGGACCCAGATTACATGGCTCTTGAAGTGTCTGATGAACCGG	840
QY	841	GATGAAAGGACCACTACGGCTGGACTGTGACTGTGGTGGTCACTGGGCGGTGATTCCTAT	900
Db	841	GATGAAAGGACCACTACGGCTGGACTGTGACTGTGGTGGTCACTGGGCGGTGATTCCTAT	900

1981	Db		AGGGAGCTGAGAAAGCTGCAGAACCGAGAGGATTTCTTCTGAAGGCAATCAGAAAAAGCTG	2040
2041	QY		GTGGAACTCAGGAAACGCGCCCAATTCCTCGAGAAACAAGGTAAAGAGACTAGAGACCATG	2100
2041	Db		GTGGAACTCAGGAAACGCGCCCAATTCCTCGAGAAACAAGGTAAAGAGACTAGAGACCATG	2100
2101	QY		GAGCGTAGAAAAACACAGACTGAAGGATGACATCCAGACAAAAATCCCAACAGATCCAGCAG	2160
2101	Db		GAGCGTAGAAAAACACAGACTGAAGGATGACATCCAGACAAAAATCCCAACAGATCCAGCAG	2160
2161	QY		ATGGCTGATAAAATTTCTGGAGCTCGAAGAGAAACATCGGAGGCGCCCAAGTCTCAGGCCAG	2220
2161	Db		ATGGCTGATAAAATTTCTGGAGCTCGAAGAGAAACATCGGAGGCGCCCAAGTCTCAGGCCAG	2220
2221	QY		CACCTGAAGTGCACCTGAAACAGAAAGACGACACTATCAGGAAAAAGATTTAAAGTGTG	2280
2221	Db		CACCTGAAGTGCACCTGAAACAGAAAGACGACACTATCAGGAAAAAGATTTAAAGTGTG	2280
2281	QY		GACAACTCAGATAAAGAAAGACCTGGCTGCAAGAGGAGACACTGCAGACATCATGATGCAGAGA	2340
2281	Db		GACAACTCAGATAAAGAAAGACCTGGCTGCAAGAGGAGACACTGCAGACATCATGATGCAGAGA	2340
2341	QY		CACGAGGAGAGCCCATGTAGAAGGGCAAAATTTCTCAGCGAAACAGAGGCGATGATCAAT	2400
2341	Db		CACGAGGAGAGCCCATGTAGAAGGGCAAAATTTCTCAGCGAAACAGAGGCGATGATCAAT	2400
2401	QY		GCTATGGATCCAAAGATCAGATCCCTGGAAACAGAGNATTCGTGNACTGTCTGAAAGCCAAT	2460
2401	Db		GCTATGGATCCAAAGATCAGATCCCTGGAAACAGAGNATTCGTGNACTGTCTGAAAGCCAAT	2460
2461	QY		AAACTTCGACCAATAGCAGTCTTTTTACCCAAAGGAAACATGAAGGCCCAAGAGAGATG	2520
2461	Db		AAACTTCGACCAATAGCAGTCTTTTTACCCAAAGGAAACATGAAGGCCCAAGAGAGATG	2520
2521	QY		ATTTCTGAACTCAGGCGAACAGAAATTTTAACTGGAGACACAGGCTGGGAAGTTGGAGGCC	2580
2521	Db		ATTTCTGAACTCAGGCGAACAGAAATTTTAACTGGAGACACAGGCTGGGAAGTTGGAGGCC	2580
2581	QY		CAGAACCGAAAACTGGAGGAGCAGCTGGAGAAAGATCAGCCACCAAGACACACAGTGACAAG	2640
2581	Db		CAGAACCGAAAACTGGAGGAGCAGCTGGAGAAAGATCAGCCACCAAGACACACAGTGACAAG	2640
2641	QY		AATCGGCTGTGAACTGGAGACAAAGATTCGGAGGTCAGTCTAGAGCACGAGGAGCAG	2700
2641	Db		AATCGGCTGTGAACTGGAGACAAAGATTCGGAGGTCAGTCTAGAGCACGAGGAGCAG	2700
2701	QY		AAACTGGAGCTCAAGCGCCAGCTCACAGACTACAGCTCTCCCTGCAGGAGCGCGAGTCA	2760
2701	Db		AAACTGGAGCTCAAGCGCCAGCTCACAGACTACAGCTCTCCCTGCAGGAGCGCGAGTCA	2760
2761	QY		CAGTTGACGCCCTGCAGGCTGCACGGCGCGCCCTGGAGAGCCAGCTTCGCCAGGCGAAG	2820
2761	Db		CAGTTGACGCCCTGCAGGCTGCACGGCGCGCCCTGGAGAGCCAGCTTCGCCAGGCGAAG	2820
2821	QY		ACAGAGCTGGAAGAGACCAACAGCAGAGCTGAAGAGGAGATCCAGGCACATCAGGCACAT	2880
2821	Db		ACAGAGCTGGAAGAGACCAACAGCAGAGCTGAAGAGGAGATCCAGGCACATCAGGCACAT	2880
2881	QY		AGAGATGAAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAAATCAGAGC	2940
2881	Db		AGAGATGAAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAAATCAGAGC	2940
2941	QY		CTGGAGGAGCAGTAAACAGAGTGCACGAGGACAAACCTGAACTCAACACCAAAAATTC	3000
2941	Db		CTGGAGGAGCAGTAAACAGAGTGCACGAGGACAAACCTGAACTCAACACCAAAAATTC	3000
3001	QY		TACTTTGTCAAAACAACTCGATGAGGCTTCTGGCGCCCAACGACGAGATTTGTACAACTCGGA	3060
3001	Db		TACTTTGTCAAAACAACTCGATGAGGCTTCTGGCGCCCAACGACGAGATTTGTACAACTCGGA	3060
3061	QY		AGTGAAGTGACCACTTCGCCCGGAGATCACCGAAACGAGAGATGACAGCTTACACGCCAG	3120

Db	3061	AGTGAAGTGGAACCATCTTCGCGCGGGAGATCACGGAAACGAGAGATGACAGCTTACACGCCAG	3120
Qy	3121	AAGCAAAACGATGAGGCGCTCTGAAGACCAACGTGCACCATGCTGGAGGAACAGGTCATCGAT	3180
Db	3121	AAGCAAAACGATGAGGCGCTCTGAAGACCAACGTGCACCATGCTGGAGGAACAGGTCATCGAT	3180
Qy	3181	TTGGAGGCCCTAAACGATGAGCTGCTTAGAAAAAGCGGCAGTGGGAGGCGCTGGAGGAGC	3240
Db	3181	TTGGAGGCCCTAAACGATGAGCTGCTTAGAAAAAGCGGCAGTGGGAGGCGCTGGAGGAGC	3240
Qy	3241	GTCTCTGGTGATCAGAGAAATCCCACTTTGAGTGTCCGGTTTCAGAGAGCTGCAGAGAAATCGTG	3300
Db	3241	GTCTCTGGTGATCAGAGAAATCCCACTTTGAGTGTCCGGTTTCAGAGAGCTGCAGAGAAATCGTG	3300
Qy	3301	GACACGAGAAAAACAGAGCAGGGCGAGGCGCATACCGGATCACCGAGTCTCGCCAGGTG	3360
Db	3301	GACACGAGAAAAACAGAGCAGGGCGAGGCGCATACCGGATCACCGAGTCTCGCCAGGTG	3360
Qy	3361	GTGGAGCTGGCAGTGAAGGAGCAAGAGCTGACATTTCTCGCTCTGCAGCAGGCTCTCAAA	3420
Db	3361	GTGGAGCTGGCAGTGAAGGAGCAAGAGCTGACATTTCTCGCTCTGCAGCAGGCTCTCAAA	3420
Qy	3421	GAGCAGAAGCTGAAGGCGGAGAGCGCTCTCTGAACAAGCTCAATGAGCTGAGAGAAAGCAT	3480
Db	3421	GAGCAGAAGCTGAAGGCGGAGAGCGCTCTCTGAACAAGCTCAATGAGCTGAGAGAAAGCAT	3480
Qy	3481	GCTATGCTTGAAATGAATGCCCGAAGCTTACACAGAGAAGCTGGAGACTGAACAGAGAGCTC	3540
Db	3481	GCTATGCTTGAAATGAATGCCCGAAGCTTACACAGAGAAGCTGGAGACTGAACAGAGAGCTC	3540
Qy	3541	AAACAGAGGCTTCTGGAAGAGCAAGCCAAATTACAGCAGCAGATGGACTCTGCAGAAAAAT	3600
Db	3541	AAACAGAGGCTTCTGGAAGAGCAAGCCAAATTACAGCAGCAGATGGACTCTGCAGAAAAAT	3600
Qy	3601	CACATTTTCGGTCTGACTCAGGACTGCAGAGAGCTCTAGATCGGCTGATCTACTGAAG	3660
Db	3601	CACATTTTCGGTCTGACTCAGGACTGCAGAGAGCTCTAGATCGGCTGATCTACTGAAG	3660
Qy	3661	ACAGAAAGAGTGACATTGGAGTATCAGCTGSEAAAAATTACAGTGTCTTATTCTCATGAA	3720
Db	3661	ACAGAAAGAGTGACATTGGAGTATCAGCTGSEAAAAATTACAGTGTCTTATTCTCATGAA	3720
Qy	3721	AAGGTGAAATGGAAGCACTATTCTCAACAAACCAACTCATGATTTCTTCGAGACC	3780
Db	3721	AAGGTGAAATGGAAGCACTATTCTCAACAAACCAACTCATGATTTCTTCGAGACC	3780
Qy	3781	AAATGGAACCACTGCTTAAAAAGAAAAGGGTTTATTTAGTCGACGGAAGAGGAGCCCT	3840
Db	3781	AAATGGAACCACTGCTTAAAAAGAAAAGGGTTTATTTAGTCGACGGAAGAGGAGCCCT	3840
Qy	3841	GCATTACCCACAACAGGTTCTCTGAGTACAATGAGCTGAAGCTGGCCCTGAGAGAGGAG	3900
Db	3841	GCATTACCCACAACAGGTTCTCTGAGTACAATGAGCTGAAGCTGGCCCTGAGAGAGGAG	3900
Qy	3901	AAAGCTCGCTGCGCAGAGCTAGAGGAGGCCCTTTCAGAAGACCCGCATCGAGCTCCGGTCC	3960
Db	3901	AAAGCTCGCTGCGCAGAGCTAGAGGAGGCCCTTTCAGAAGACCCGCATCGAGCTCCGGTCC	3960
Qy	3961	GCCCGGAGGAAGCTGCCACCGCAAGCAACCGGACCAACCCATCCACGCCAGCC	4020
Db	3961	GCCCGGAGGAAGCTGCCACCGCAAGCAACCGGACCAACCCATCCACGCCAGCC	4020
Qy	4021	ACCGGAGGCAGAGATCGCATGCTCGCCATCGTGGCTGCGCCAGAGCACCGCCAGT	4080
Db	4021	ACCGGAGGCAGAGATCGCATGCTCGCCATCGTGGCTGCGCCAGAGCACCGCCAGT	4080
Qy	4081	GCCATGAGCCTGTGGCGCCGCCCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAA	4140
Db	4081	GCCATGAGCCTGTGGCGCCGCCCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAA	4140
Qy	4141	TTTAGTCGGCGCTCTTAAGGAACCGATGCACCAACAATATTCCTCACCGATTCAGCTAGGA	4200
Db	4141	TTTAGTCGGCGCTCTTAAGGAACCGATGCACCAACAATATTCCTCACCGATTCAGCTAGGA	4200





Db 1501 TCCCTCTGGAGCAGGACCTTGTACTTACATCACAGATCCAGTAGCTTAAGCGAAGT 1560  
Qy 1579 TTGAGCAAGCAGCATGGAGGTGTCCAGGAGATGACAAAGCACTGCGAGCTTCTCCAT 1638  
Db 1561 TTGGAGCAAGCAGCATGGAGGTGTCCAGGAGATGACAAAGCACTGCGAGCTTCTCCAT 1620  
Qy 1639 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGCAGGAGTACCGAGGTCAA 1698  
Db 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGCAGGAGTACCGAGGTCAA 1680  
Qy 1699 GTGGAAGAAATGAGGTGATGATGAATCAATGATGGAAGAGGATCTTGTCTCAGCAAGAAGA 1758  
Db 1681 GTGGAAGAAATGAGGTGATGATGAATCAATGATGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740  
Qy 1759 CGGAGTGAATCTACGATCTGAGTCTGAGAGAGTCTCGGCTTCTGCTGGAAGAAATTCAG 1818  
Db 1741 CGGAGTGAATCTACGATCTGAGTCTGAGAGAGTCTCGGCTTCTGCTGGAAGAAATTCAG 1800  
Qy 1819 CGGAAAGCGACAGAAATGTTCAGCATAAACTGTGTAAGGCTTAAGGATCAAGGAAAGCCTGAA 1878  
Db 1801 CGGAAAGCGACAGAAATGTTCAGCATAAACTGTGTAAGGCTTAAGGATCAAGGAAAGCCTGAA 1860  
Qy 1879 GTGGAGAAATATGGAAGCTGGAGAAATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938  
Db 1861 GTGGAGAAATATGGAAGCTGGAGAAATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
Qy 1939 CTCGAAGAGAAATCTGGAGAAAGGCTGCAAGAGCGAGCCGAGAGGAGCTGGAGAAAGCTG 1998  
Db 1921 CTCGAAGAGAAATCTGGAGAAAGGCTGCAAGAGCGAGCCGAGAGGAGCTGGAGAAAGCTG 1980  
Qy 1999 CAGAACCGAGAGGATTTCTTGAAGGATCAGAAAGAAAGCTGTGTGAAGCTGAGGAACGC 2058  
Db 1981 CAGAACCGAGAGGATTTCTTGAAGGATCAGAAAGAAAGCTGTGTGAAGCTGAGGAACGC 2040  
Qy 2059 CGCCATTTCTTGGAGAACAAAGTAAAGAGACTAGAGACCATGAGCGTAGAGAAACAGA 2118  
Db 2041 CGCCATTTCTTGGAGAACAAAGTAAAGAGACTAGAGACCATGAGCGTAGAGAAACAGA 2100  
Qy 2119 CTGAAGATGACATCCAGACAAATCCAAACAGATCCAGCAGATGGCTGATAAATTTCTG 2178  
Db 2101 CTGAAGATGACATCCAGACAAATCCAAACAGATCCAGCAGATGGCTGATAAATTTCTG 2160  
Qy 2179 GAGCTCGAAGAAACATCGGGAGGCCAGTCTCAGCCAGCAGCAGTACAGTGCACCTG 2238  
Db 2161 GAGCTCGAAGAAACATCGGGAGGCCAGTCTCAGCCAGCAGCAGTACAGTGCACCTG 2220  
Qy 2239 AAACAGAAAGAGCAGCAGCTATGAGGAAAGATTAAGTGTGAGCAATCAGATAAAGAAA 2298  
Db 2221 AAACAGAAAGAGCAGCAGCTATGAGGAAAGATTAAGTGTGAGCAATCAGATAAAGAAA 2280  
Qy 2299 GAGCTGGCTGACAGAGACACTGGAGAACATGATCCAGAGACACAGAGAGAGGCCCAT 2358  
Db 2281 GAGCTGGCTGACAGAGACACTGGAGAACATGATCCAGAGACACAGAGAGAGGCCCAT 2340  
Qy 2359 GAGAAAGGCAAAATTTCTCAGCGAACAGAGGCGATGATCAATGCTATGATTTCCAGATC 2418  
Db 2341 GAGAAAGGCAAAATTTCTCAGCGAACAGAGGCGATGATCAATGCTATGATTTCCAGATC 2400  
Qy 2419 AGATCCCTGGAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTCAGCAAAATAGC 2478  
Db 2401 AGATCCCTGGAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTCAGCAAAATAGC 2460  
Qy 2479 AGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAA 2538  
Db 2461 AGTCTTTTACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAA 2520  
Qy 2539 CAGAAATTTTACCTGAGACACAGGCTGGAAAGTTGGAGGCCCAAGAACCGAAATTCGGAG 2598  
Db 2521 CAGAAATTTTACCTGAGACACAGGCTGGAAAGTTGGAGGCCCAAGAACCGAAATTCGGAG 2580  
Qy 2599 GAGCAGCTGGAGAGATCAGCCACCAAGACACAGTGAAGAAATCGGCTGTGGAACCTG 2658  
Db 2581 GAGCAGCTGGAGAGATCAGCCACCAAGACACAGTGAAGAAATTCGGCTGTGGAACCTG 2640

Qy 2659 GAGACAAAGATTCGGGAGAGTCACTTCTAGAGCACGAGGAGCAAACTGGAGCTCAAGCGC 2718  
Db 2641 GAGACAAAGATTCGGGAGAGTCACTTCTAGAGCACGAGGAGCAAACTGGAGCTCAAGCGC 2700  
Qy 2719 CAGCTCAGAGACTACAGCTCTCCCTGAGAGCGCGAGTCAAGTTTACAGCCCTGAG 2778  
Db 2701 CAGCTCAGAGACTACAGCTCTCCCTGAGAGCGCGAGTCAAGTTTACAGCCCTGAG 2760  
Qy 2779 GCTGACGGGGGGCCCTGGAGAGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACC 2838  
Db 2761 GCTGACGGGGGGCCCTGGAGAGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACC 2820  
Qy 2839 ACAGCAGAAAGCTGAAGAGAGATCCAGCACTCAGGCAATAGAGATGAATTCAGGCGC 2898  
Db 2821 ACAGCAGAAAGCTGAAGAGAGATCCAGCACTCAGGCAATAGAGATGAATTCAGGCGC 2880  
Qy 2899 AAATTTGATGCTCTTCGTAAACAGAGTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC 2958  
Db 2881 AAATTTGATGCTCTTCGTAAACAGAGTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC 2940  
Qy 2959 CAGCTGACCGAGGACACGCTGAACTCAAACCAAACTTCTACTTGTCCAAACAACTC 3018  
Db 2941 CAGCTGACCGAGGACACGCTGAACTCAAACCAAACTTCTACTTGTCCAAACAACTC 3000  
Qy 3019 GATGAGGCTTCTGGCGCCCAACGAGAGATTGTACAACTGCGAAGTGAAGTGGACCATCTC 3078  
Db 3001 GATGAGGCTTCTGGCGCCCAACGAGAGATTGTACAACTGCGAAGTGAAGTGGACCATCTC 3060  
Qy 3079 CGCCGGAGATCAACGAAACGAGAGATCAGCTTACCAGCCAGAAACCAAGTGAAGGCT 3138  
Db 3061 CGCCGGAGATCAACGAAACGAGAGATCAGCTTACCAGCCAGAAACCAAGTGAAGGCT 3120  
Qy 3139 CTGAAGACCACTGCACCATCTCGAGGAAACAGCTCATGATTTGGAGGCCCTTAAACGAT 3198  
Db 3121 CTGAAGACCACTGCACCATCTCGAGGAAACAGCTCATGATTTGGAGGCCCTTAAACGAT 3180  
Qy 3199 GAGCTGCTAGAAAAAGAGCGGAGCTGGAGGCGCTGGAGGAGCGCTCTGGGTGATGAGAAA 3258  
Db 3181 GAGCTGCTAGAAAAAGAGCGGAGCTGGAGGCGCTGGAGGAGCGCTCTGGGTGATGAGAAA 3240  
Qy 3259 TCCAGTTTGAGTGTCCGGTTCGAGAGCTGAGAGANTGCTGCACCCAGAGAACAGAGC 3318  
Db 3241 TCCAGTTTGAGTGTCCGGTTCGAGAGCTGAGAGANTGCTGCACCCAGAGAACAGAGC 3300  
Qy 3319 AGGCGAGAGCCGATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG 3378  
Db 3301 AGGCGAGAGCCGATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG 3360  
Qy 3379 GAGCAGAGGCTGAGATTTCTGCTCTCAGAGGCTCTCAAAGAGCAGAGCTGAAGGCC 3438  
Db 3361 GAGCAGAGGCTGAGATTTCTGCTCTCAGAGGCTCTCAAAGAGCAGAGCTGAAGGCC 3420  
Qy 3439 GAGAGCTCTCTGACAAAGCTCAATGACCTGAGAAAGCATGCTATGCTTGAATGAAT 3498  
Db 3421 GAGAGCTCTCTGACAAAGCTCAATGACCTGAGAAAGCATGCTATGCTTGAATGAAT 3480  
Qy 3499 GCCGAAAGCTTACAGCAGAGCTGGAGACTGAACAGAGCTCAAAACAGAGGCTTCTGGAA 3558  
Db 3481 GCCGAAAGCTTACAGCAGAGCTGGAGACTGAACAGAGCTCAAAACAGAGGCTTCTGGAA 3540  
Qy 3559 GAGCAGCCAAATTTACAGCAGCAGAGTGGACCTCGAGAAATTCACATTTTCCGCTGACT 3618  
Db 3541 GAGCAGCCAAATTTACAGCAGCAGAGTGGACCTCGAGAAATTCACATTTTCCGCTGACT 3600  
Qy 3619 CAAGGACTGCAAGAGCTCTAGATCGGGTGTACTACTGAAGACAGAAAGTGAAGCTTG 3678  
Db 3601 CAAGGACTGCAAGAGCTCTAGATCGGGTGTACTACTGAAGACAGAAAGTGAAGCTTG 3660  
Qy 3679 GAGTATCAGCTGGAAGAACTTCAAGTCTCTATTTCTCATGAAGGCTGAAGTGGAGGC 3738  
Db 3661 GAGTATCAGCTGGAAGAACTTCAAGTCTCTATTTCTCATGAAGGCTGAAGTGGAGGC 3720

QY	3739	ACTATTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGCCAAATATGACCAACTGCT	3798
Db	3721	ACTATTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGCCAAATATGACCAACTGCT	3780
QY	3799	AAAAAGAAAAGGGTTATTATTAGTCGACGGAAGAGGCCCTGCTTTTACCCACACAGGTT	3858
Db	3781	AAAAAGAAAAGGGTTATTATTAGTCGACGGAAGAGGCCCTGCTTTTACCCACACAGGTT	3840
QY	3859	CCTCTCAGTACAAATGAGCTGAAGCTGGCCCTGGAGAAGGAAAGCTCGCTGTGCAGAG	3918
Db	3841	CCTCTCAGTACAAATGAGCTGAAGCTGGCCCTGGAGAAGGAAAGCTCGCTGTGCAGAG	3900
QY	3919	CTAGAGGAAGCCCTTCAGAGAGCCGCATCGAGCTCCGCTCCGCCCGGAGGAGAGTGCC	3978
Db	3901	CTAGAGGAAGCCCTTCAGAGAGCCGCATCGAGCTCCGCTCCGCCCGGAGGAGAGTGCC	3960
QY	3979	CACGCAAAACGACGAGCAACCCACACCCATCCACGCCAGCCACCCGCGAGGACGACATC	4038
Db	3961	CACGCAAAACGACGAGCAACCCACACCCATCCACGCCAGCCACCCGCGAGGACGACATC	4020
QY	4039	GCCATGTCGCCCATCTGTGGGTGCGCAGAGCACACGCCAGCCAGTCGCATAGACCTGCTGGCC	4098
Db	4021	GCCATGTCGCCCATCTGTGGGTGCGCAGAGCACACGCCAGCCAGTCGCATAGACCTGCTGGCC	4080
QY	4099	CCGCCATCCAGCCGCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAG	4158
Db	4081	CCGCCATCCAGCCGCGAGAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAG	4140
QY	4159	GAACGCAATGCACACAATATCTCTACCGGATTCAAAGTAGGACTGAACATGCGAGCCACA	4218
Db	4141	GAACGCAATGCACACAATATCTCTACCGGATTCAAAGTAGGACTGAACATGCGAGCCACA	4200
QY	4219	AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGAGCCAGGCACTCAAATCTCTCGAA	4278
Db	4201	AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGAGCCAGGCACTCAAATCTCTCGAA	4260
QY	4279	TGTCAGGTGATGTCTCACCCCAAAGTGCTCCACGTGTTGCCAGCCACTGCGGCTTGCCCT	4338
Db	4261	TGTCAGGTGATGTCTCACCCCAAAGTGCTCCACGTGTTGCCAGCCACTGCGGCTTGCCCT	4320
QY	4339	GCTGTAATATGCCACACATTCACCGAGGCGCTTCTGCCGTGACAAATGAACCTCCACAGT	4398
Db	4321	GCTGTAATATGCCACACATTCACCGAGGCGCTTCTGCCGTGACAAATGAACCTCCACAGT	4380
QY	4399	CTCCAGAACCAAGAGCCCGACGACAGCTTGCACTCTGGAAGGGTGGATGAAGTGCCCAAG	4458
Db	4381	CTCCAGAACCAAGAGCCCGACGACAGCTTGCACTCTGGAAGGGTGGATGAAGTGCCCAAG	4440
QY	4459	AATAAACCAACGAGGACAGCAAGGCTGGGACAGGAAGTACATTTGCTGGAGGGATCAAAA	4518
Db	4441	AATAAACCAACGAGGACAGCAAGGCTGGGACAGGAAGTACATTTGCTGGAGGGATCAAAA	4500
QY	4519	GTCTCTATTATGACAAATGAAGCCAGAGAAGCTGGACAGAGGCGGTGGGAAGAATTTGAG	4578
Db	4501	GTCTCTATTATGACAAATGAAGCCAGAGAAGCTGGACAGAGGCGGTGGGAAGAATTTGAG	4560
QY	4579	CTGTGCCCTCCCGACGGGATGTATCTATTATCATGTGTCCGTTGGTGCTTCGGAACCTGCA	4638
Db	4561	CTGTGCCCTCCCGACGGGATGTATCTATTATCATGTGTCCGTTGGTGCTTCGGAACCTGCA	4620
QY	4639	AATACGCCAAAGCAGAAAAAGCAGAGAAGCTGATCTAAATGCTTGAAACTCCCTGCTG	4698
Db	4621	AATACGCCAAAGCAGAAAAAGCAGAGAAGCTGATCTAAATGCTTGAAACTCCCTGCTG	4680
QY	4699	AAACTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGCTGCGCTTCACTGACCAAGTG	4758
Db	4681	AAACTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGCTGCGCTTCACTGACCAAGTG	4740
QY	4759	GTGTTGCTGGCACCGGAGGAGGCTCTACGCCCTGGAATGCTTTGAAAAACTCCCTAACC	4818
Db	4741	GTGTTGCTGGCACCGGAGGAGGCTCTACGCCCTGGAATGCTTTGAAAAACTCCCTAACC	4800
QY	4819	CATGTCGCCAGGAATGGAGCAGTCTTCCAAATTTATATTATCAAGGACCTGGGAAGCTA	4878

4801	Db	CATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATCAAGGACCTGGAGAAGCTTA4860
4879	QY	CTCATGATAGCAGGAAGAGCGGGCACTGTGTCTTTTGGACGTGAAGAAAGTGAACAG4938
4861	Db	CTCATGATAGCAGGAAGAGCGGGCACTGTGTCTTTGGACGTGAAGAAAGTGAACAG4920
4939	QY	TCCTTGGCCAGTCCCACTGCTGCCCCAGCCGACATCTCAGCCCAACATTTTITGAAGCT4998
4921	Db	TCCTTGGCCAGTCCCACTGCTGCCCCAGCCGACATCTCAGCCCAACATTTTITGAAGCT4980
4999	QY	GTCAAGGGCTGCCACATTTGTTGGGGCAGGCAAGATTGAGAACGGGCTCTGCATCTGTGCA5058
4981	Db	GTCAAGGGCTGCCACTTTGTTGGGGCAGGCAAGATTGAGAACGGGCTCTGCATCTGTGCA5040
5059	QY	GCCATGCCAGCAAGTGTCTCATTTCTCCGTACAAAGAAACCTCAGCAAAATCTGCATC5118
5041	Db	GCCATGCCAGCAAGTGTCTCATTTCTCCGTACAAAGAAACCTCAGCAAAATCTGCATC5100
5119	QY	CGGAAAGAGATAGACCTCAGAGCCCTGCAGCTGTATCCACTTCACCAATTCAGATATC5178
5101	Db	CGGAAAGAGATAGACCTCAGAGCCCTGCAGCTGTATCCACTTCACCAATTCAGATATC5160
5179	QY	CTCATTTGGAACCAATTAATTTCTACGAAATTCGATGAGACGATACAGCTTCGAGGAAATTC5238
5161	Db	CTCATTTGGAACCAATTAATTTCTACGAAATTCGATGAGACGATACAGCTTCGAGGAAATTC5220
5239	QY	CTCGATAAGAAATGACCAATTCCTTTGGACCTGCTGTGTTTGCCTGCTCTTCCACACAGCTTC5298
5221	Db	CTCGATAAGAAATGACCAATTCCTTTGGACCTGCTGTGTTTGCCTGCTCTTCCACACAGCTTC5280
5299	QY	CCTGTCTCAATCTGTGAGTGTAACAGCGCAGGCGCAGAGGAGTGACTTTCGTGTGTTTC5358
5281	Db	CCTGTCTCAATCTGTGAGTGTAACAGCGCAGGCGCAGAGGAGTGACTTTCGTGTGTTTC5340
5359	QY	CAGCAATTTGGAGTGTCGTGGATTCTTACGGAAGACGTAGCCGCAAGACGATCTCAAG5418
5341	Db	CAGCAATTTGGAGTGTCGTGGATTCTTACGGAAGACGTAGCCGCAAGACGATCTCAAG5400
5419	QY	TGGAGTGCCTTACCTTTGGCTTTGGCTACAGAGAACCTTATCTGTTTGTGACCCACTTC5478
5401	Db	TGGAGTGCCTTACCTTTGGCTTTGGCTACAGAGAACCTTATCTGTTTGTGACCCACTTC5460
5479	QY	AATCTACTCGAAGTAATTGATATCCAGGCACGTCTCTCAGCAGGGACCCCTGCCCGAGCG5538
5461	Db	AATCTACTCGAAGTAATTGATATCCAGGCACGTCTCTCAGCAGGGACCCCTGCCCGAGCG5520
5539	QY	TACTCTGACATCCGAAACCCGCGCTACCTGGGCCCCCTGCCATTTCTCTCAGGAGCGAATTAC5598
5521	Db	TACTCTGACATCCGAAACCCGCGCTACCTGGGCCCCCTGCCATTTCTCTCAGGAGCGAATTAC5580
5599	QY	TTGGCGTCTCTATACGAGGATAAATTAAAGGTCAATTGCTGCAAGGAACCTCGTGAAG5658
5581	Db	TTGGCGTCTCTATACGAGGATAAATTAAAGGTCAATTGCTGCAAGGAACCTCGTGAAG5640
5659	QY	GAGTCCGGCACTGAAACACCCGCGGGCCCTGCACCTCCCGCAGCAGCCCAACAAGCGGA5718
5641	Db	GAGTCCGGCACTGAAACACCCGCGGGCCCTGCACCTCCCGCAGCAGCCCAACAAGCGGA5700
5719	QY	GGCCCCACCGTACAAAGAGCAATCAACAAAGCGGTGGCTTCACGCCAGCGCCCGCC5778
5701	Db	GGCCCCACCGTACAAAGAGCAATCAACAAAGCGGTGGCTTCACGCCAGCGCCCGCC5760
5779	QY	GAGGGCCCCAGCCACCGCAGAGCCCAAGCACACCCACCGCTACCCGAGGGCGCGAC5838
5761	Db	GAGGGCCCCAGCCACCGCAGAGCCCAAGCACACCCACCGCTACCCGAGGGCGCGAC5820
5839	QY	GAGCTGCGCAGGGACAAGTCTCTTGGCGCCGCCCTGTGAGGAGAGAGTCTCCCGCGCGG5898
5821	Db	GAGCTGCGCAGGGACAAGTCTCTTGGCGCCGCCCTGTGAGGAGAGAGTCTCCCGCGCGG5880
5899	QY	ATGCTCAGCACGCGGAGAGACCGTCCCCCGGGAGGCTGTTTGAAGACAGCAGCAGGAGG5958



QY 1441 TTACATCGGAGAGTGTACAGAGTGGAGGCTGTGCTTAGTTCAGAAAGAGGTGGAGCTGAAG 1500  
DB 1477 TTACATCGGAGAGTGTACAGAGTGGAGGCTGTGCTTAGTTCAGAAAGAGGTGGAGCTGAAG 1536  
QY 1501 GCTCTGAGACTCAGAGATGCCCTCTCTGGAGCAGGACCTTGTCTACCTACATCAGAAATGC 1560  
DB 1537 GCCTCTGAGACTCAGAGATGCCCTCTCTGGAGCAGGACCTTGTCTACCTACATCAGAAATGC 1596  
QY 1561 AGTAGCTTAAGCGAAGTTTGAGCAAGCAGCAGGATGGAGGTCTCCAGAGGATCAGAAA 1620  
DB 1597 AGTAGCTTAAGCGAAGTTTGAGCAAGCAGCAGGATGGAGGTCTCCAGAGGATCAGAAA 1656  
QY 1621 GCACTGCAGCTTCTCCATGATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAG 1680  
DB 1657 GCACTGCAGCTTCTCCATGATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAG 1716  
QY 1681 CAGGAGTACAGGCTCAAGTGAAGAAATGAGGTTGATGATCAATCAGTTGGAAGAGAT 1740  
DB 1717 CAGGAGTACAGGCTCAAGTGAAGAAATGAGGTTGATGATCAATCAGTTGGAAGAGAT 1776  
QY 1741 CTTTGTCTCAGCAAGAACGGAAGTGTATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTT 1800  
DB 1777 CTTTGTCTCAGCAAGAACGGAAGTGTATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTT 1836  
QY 1801 GTGTCTGAAGAAATTAAGCGGAAAGCGACAGAAATGTACAGATAAATCTGTTGAAGCTTAAG 1860  
DB 1837 GTGTCTGAAGAAATTAAGCGGAAAGCGACAGAAATGTACAGATAAATCTGTTGAAGCTTAAG 1896  
QY 1861 GATCAAGGGAAGCCTGAAGTGGAGAAATATCGAAATCTGGAAGATCAATCTCAGCAG 1920  
DB 1897 GATCAAGGGAAGCCTGAAGTGGAGAAATATCGAAATCTGGAAGATCAATCTCAGCAG 1956  
QY 1921 CAGCTCAAAATTCAGGAGCTCCAAGAGAAATCTGGAGAGGCT- 1962  
DB 1957 CAGCTCAAAATTCAGGAGCTCCAAGAGAAATCTGGAGAGGCTGTAAGAACCCAGCAGCGAG 2016  
QY 1963 -----GCAAGAGCGAGCGAGGAGGCTGGAG 1992  
DB 2017 GGCACCGAGCTGTGCAGAAATTCGCCAGGCAAGAGGAGCGAGCGGAGGCTGGAG 2076  
QY 1993 AAGCTGCAGAACCGAGAGGATCTTCTGAAGGCATCAGAAAGAAAGCTGGTGAAGCTGAG 2052  
DB 2077 AAGCTGCAGAACCGAGAGGATCTTCTGAAGGCATCAGAAAGAAAGCTGGTGAAGCTGAG 2136  
QY 2053 GAAACGGCGCATCTCTGAGAAACAAGGTAAGAGACTAGAGACCAATGAGGCGGTAGAGAA 2112  
DB 2137 GAAACGGCGCATCTCTGAGAAACAAGGTAAGAGACTAGAGACCAATGAGGCGGTAGAGAA 2196  
QY 2113 AACAGACTGAAGGATCACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAA 2172  
DB 2197 AACAGACTGAAGGATCACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAA 2256  
QY 2173 ATTCTGGAGCTCGAAGAGAAACATCCGGAGGCCCAAGTCTCAGCCCAAGCACCTAGAAGTG 2232  
DB 2257 ATTCTGGAGCTCGAAGAGAAACATCCGGAGGCCCAAGTCTCAGCCCAAGCACCTAGAAGTG 2316  
QY 2233 CAGCTGAAACAGAAAGCAGACTATGAGGAAAGATTAAGTGTGACACATCAGATA 2292  
DB 2317 CAGCTGAAACAGAAAGCAGACTATGAGGAAAGATTAAGTGTGACACATCAGATA 2376  
QY 2293 AAGAAAGACCTGGCTGACAGGAGACACTGGAGAACATGATGTCAGAGACACAGAGAGGAG 2352  
DB 2377 AAGAAAGACCTGGCTGACAGGAGACACTGGAGAACATGATGTCAGAGACACAGAGAGGAG 2436  
QY 2353 GCCCATGAGAGGGGAAAATCTCAGGAAACAGAAAGCGCATGATCAATGCTATGATTC 2412  
DB 2437 GCCCATGAGAGGGGAAAATCTCAGGAAACAGAAAGCGCATGATCAATGCTATGATTC 2496  
QY 2413 AAGATCAGATCCCTGGAAACAGAGGATTTGTGGAATCTGTGAAGCCCAATAACTTGCAGCA 2472  
DB 2497 AAGATCAGATCCCTGGAAACAGAGGATTTGTGGAATCTGTGAAGCCCAATAACTTGCAGCA 2556

QY 2473 AATAGCAGTCTTTTACCCAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATC 2532  
DB 2557 AATAGCAGTCTTTTACCCAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATC 2616  
QY 2533 AGGCAACAGAAAATTTTACCTGGAGACACAGCTGGGAAGTTGGAGGCCCAAGAACCGAAA 2592  
DB 2617 AGGCAACAGAAAATTTTACCTGGAGACACAGCTGGGAAGTTGGAGGCCCAAGAACCGAAA 2676  
QY 2593 CTGAGGAGCAGCTGGAGAAAGATCAGCCACCAAGACCACAGTGCACAGAAATCGCTGCTG 2652  
DB 2677 CTGAGGAGCAGCTGGAGAAAGATCAGCCACCAAGACCACAGTGCACAGAAATCGCTGCTG 2736  
QY 2653 GAATCTGAGACAGATTGCGGAGGTGAGTCTTAGAGCACGAGGAGCAGAACTGGAGCTC 2712  
DB 2737 GAATCTGAGACAGATTGCGGAGGTGAGTCTTAGAGCACGAGGAGCAGAACTGGAGCTC 2796  
QY 2713 AAGCGCCAGCTCAGAGAGCTACAGTCTCTCCTTGAGGAGCGCAGTGCACAGTTGACAGCC 2772  
DB 2797 AAGCGCCAGCTCAGAGAGCTACAGTCTCTCCTTGAGGAGCGCAGTGCACAGTTGACAGCC 2856  
QY 2773 CTGAGGCTGACGCGCGGCTGGAGAGCCAGCTTCGCGAGCGGAAAGACAGAGCTGGAA 2832  
DB 2857 CTGAGGCTGACGCGCGGCTGGAGAGCCAGCTTCGCGAGCGGAAAGACAGAGCTGGAA 2916  
QY 2833 GAGACACAGCAGAAAGCTGAAGAGGAGATCCAGGCACTCAGGCACATAGAGATGAATC 2892  
DB 2917 GAGACACAGCAGAAAGCTGAAGAGGAGATCCAGGCACTCAGGCACATAGAGATGAATC 2976  
QY 2893 CAGCGCAAAATTTGATGCTCTTCTGTAAACAGCTGTACTGTAAATCAGAGACCTGGAGAGCAG 2952  
DB 2977 CAGCGCAAAATTTGATGCTCTTCTGTAAACAGCTGTACTGTAAATCAGAGACCTGGAGAGCAG 3036  
QY 2953 CTAACACAGCTGACCGGAGGACAGCTGAACTCAACCAACCAAACTTCTACTTGTCCAAA 3012  
DB 3037 CTAACACAGCTGACCGGAGGACAGCTGAACTCAACCAACCAAACTTCTACTTGTCCAAA 3096  
QY 3013 CAATCTCAGTGAAGCTTCTGCGCCCAACGAGAGATTTACAACCTGCGAAAGTGAAGTGAC 3072  
DB 3097 CAATCTCAGTGAAGCTTCTGCGCCCAACGAGAGATTTACAACCTGCGAAAGTGAAGTGAC 3156  
QY 3073 CATCTCCGCGGAGATCAGGAAACAGAGATGCGCTTACCAGCCAGAGACCAACGATG 3132  
DB 3157 CATCTCCGCGGAGATCAGGAAACAGAGATGCGCTTACCAGCCAGAGACCAACGATG 3216  
QY 3133 GAGCTCTGAGACACAGCTGCACCATCTGGAGGAAACAGGTCATGGAATTTGGAGGCCCTA 3192  
DB 3217 GAGCTCTGAGACACAGCTGCACCATCTGGAGGAAACAGGTCATGGAATTTGGAGGCCCTA 3276  
QY 3193 AACGATGAGCTGTAGAAAAGAGCGCAGTGGAGGCTGGAGAGCGCTCTGGGTGAT 3252  
DB 3277 AACGATGAGCTGTAGAAAAGAGCGCAGTGGAGGCTGGAGAGCGCTCTGGGTGAT 3336  
QY 3253 GAGAAATCCAGTTTGAGTGTGCGGTTGAGAGCTGCAGAGAAATGCTGGACACCGAGAA 3312  
DB 3337 GAGAAATCCAGTTTGAGTGTGCGGTTGAGAGCTGCAGAGAAATGCTGGACACCGAGAA 3396  
QY 3313 CAGAGCAGGCGGAGAGCGGATCAGCGGATCACCGAGTCTGCCAGGTGTTGGAGCTGGCA 3372  
DB 3397 CAGAGCAGGCGGAGAGCGGATCAGCGGATCACCGAGTCTGCCAGGTGTTGGAGCTGGCA 3456  
QY 3373 GTGAAGAGACAAAGGCTGAGATTTCTGCTCTGAGCAGGCTCTCAAGAGCAGAACTG 3432  
DB 3457 GTGAAGAGACAAAGGCTGAGATTTCTGCTCTGAGCAGGCTCTCAAGAGCAGAACTG 3516  
QY 3433 AAGCCGAGAGCTCTCTGCAAGCTCAATGACCTGGAGAAAGAGCATGCTATCTTGA 3492  
DB 3517 AAGCCGAGAGCTCTCTGCAAGCTCAATGACCTGGAGAAAGAGCATGCTATCTTGA 3576  
QY 3493 ATGAATGCCGAAGCTTACAGCAGAAAGCTGGAGACTGAAACAGAGCTCAACAGAGGCTT 3552  
DB 3577 ATGAATGCCGAAGCTTACAGCAGAAAGCTGGAGACTGAAACAGAGCTCAACAGAGGCTT 3636  
QY 3553 CTGGAAGAGCAGCCAAATTTACACGACAGATGGACCTGCAGAAAAATCATTCTTCGT 3612



```
QY 5623 TTAAGGGTCATTTGCTGCAAGGGAACCTCTGTGAAGGAGTCCGGCACTGAACACACCACCGG 5682
Db 5812 TTAAGGGTCATTTGCTCAAGGGAACCTCTGTGAAGGAGTCCGGCACTGAACACACCACCGG 5871
QY 5683 GGCCTGTCACCTCCCGCAGCAGCCCAACCAAGCAGGAGCCCAACCAAGCAGGAGCAGC 5742
Db 5872 GGCCTGTCACCTCCCGCAGCAGCCCAACCAAGCAGGAGCCCAACCAAGCAGGAGCAGC 5931
QY 5743 ATCAACAAGCGGTGGCTTCCAGCCAGCGCCGCGCCGAAAGCCCAACCAAGCAGGAG 5802
Db 5932 ATCAACAAGCGGTGGCTTCCAGCCAGCGCCGCGCCGAAAGCCCAACCAAGCAGGAG 5991
QY 5803 CCAAGCAGCACCACCGCTACCGCAGGAGGCGGAGCCGAGCTGCGCAGGAGCAGGCTCCT 5862
Db 5992 CCAAGCAGCACCACCGCTACCGCAGGAGGCGGAGCCGAGCTGCGCAGGAGCAGGCTCCT 6051
QY 5863 GGCCTGTCACCTCCCGCAGCAGCCCAACCAAGCAGGAGCCCAACCAAGCAGGAGCAGC 5922
Db 6052 GGCCTGTCACCTCCCGCAGCAGCCCAACCAAGCAGGAGCCCAACCAAGCAGGAGCAGC 6111
QY 5923 TCCCGCGGAGCTGTTGAAGCAGCAGCAGGAGGCGCGCTGCTGCGGAGCAGGAGGAGG 5982
Db 6112 TCCCGCGGAGCTGTTGAAGCAGCAGCAGGAGGCGCGCTGCTGCGGAGCAGGAGGAGG 6171
QY 5983 ACCCGCTGTCAGGAGTGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6042
Db 6172 ACCCGCTGTCAGGAGTGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6231
QY 6043 AACCTGTCACCTATTAT 6060
Db 6232 AAAACCAACTCCTCATCT 6249

RESULT 4
AX6711108 8603 bp DNA linear PAT 27-MAR-2003
LOCUS Sequence 4 from Patent WO03004523.
DEFINITION AX6711108
ACCESSION AX6711108
VERSION AX6711108.1 GI:29329570
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Zhu, Z.
TITLE Regulation of human citron rho/rac-interacting kinase
JOURNAL Patent: WO 03004523-A 4 16-JAN-2003;
Bayer Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
1..8603
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 86.2%; Score 5668; DB 6; Length 8603;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 5962; Conservative 0; Mismatches 35; Indels 243; Gaps 3;

QY 19 ATGTTGAAGTCAAAATATGAGCGGGGAATCCTTTGGATGCTGGTCTGAAACCCATT 78
Db 1 ATGTTGAAGTCAAAATATGAGCGGGGAATCCTTTGGATGCTGGTCTGAAACCCATT 60
QY 79 GCCAGCGGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACCCACCTTTATGACTCAA 138
Db 61 GCCAGCGGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACCCACCTTTATGACTCAA 120
QY 139 CAGCAGATGTCCTCTTTCCGAGAGGAGATATAGATGCCCTCTTTGTTCTCTTTGAA 198
Db 121 CAGCAGATGTCCTCTTTCCGAGAGGAGATATAGATGCCCTCTTTGTTCTCTTTGAA 180
```

```
QY 199 GAATGAGTCAGCCTGCTCTGATGAAGATTAAAGCAGTGAAGCACTTTGTCGGAGTAT 258
Db 181 GAATGAGTCAGCCTGCTCTGATGAAGATTAAAGCAGTGAAGCACTTTGTCGGAGTAT 240
QY 259 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAAGTCAGA 318
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAAGTCAGA 300
QY 319 AGTCTTGAGTGTGGTCACTTTGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 378
Db 301 AGTCTTGAGTGTGGTCACTTTGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
QY 379 GACATCTATCTTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 438
Db 361 GACATCTATCTTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
QY 439 TTTTTCGAGAAAGCGGAAACATATTATCTGAAGCACAAGCCCGTGGATTCCTCCCAATTA 498
Db 421 TTTTTCGAGAAAGCGGAAACATATTATCTGAAGCACAAGCCCGTGGATTCCTCCCAATTA 480
QY 499 CAGTATGCTTTCAGGACAAAATCCTTTATCTGATGAGGAATATCAGCTGGAGGG 558
Db 481 CAGTATGCTTTCAGGACAAAATCCTTTATCTGATGAGGAATATCAGCTGGAGGG 540
QY 559 GACTTGCTGTCACCTTTTGAATAGATATGAGGACCAAGTTAGATGAAAACCTGATACAGTTT 618
Db 541 GACTTGCTGTCACCTTTTGAATAGATATGAGGACCAAGTTAGATGAAAACCTGATACAGTTT 600
QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGGTTTCAAGGTTTCAAGTGAAGTGAAGTGAAG 678
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGTTTCAAGGTTTCAAGTGAAGTGAAGTGAAG 660
QY 679 GACATCAAGCTGAGAAACATTTCTGTTGACCGCAGCAGGACACATCAAGCTGGTGGATTTT 738
Db 661 GACATCAAGCTGAGAAACATTTCTGTTGACCGCAGCAGGACACATCAAGCTGGTGGATTTT 720
QY 739 GGATCTCCGCGAAAATGAATTCAAAACAGATGCTGAATGCCAAAACCTCCCGATGGAGCC 798
Db 721 GGATCTCCGCGAAAATGAATTCAAAACAGATGCTGAATGCCAAAACCTCCCGATGGAGCC 780
QY 799 CCAGATTACATGCTCCTGAAAGTGTGACTGTGATGAACGGGATGGAAGGACCACTAC 858
Db 781 CCAGATTACATGCTCCTGAAAGTGTGACTGTGATGAACGGGATGGAAGGACCACTAC 840
QY 859 GGCCTGAGCTGTGACTGTGGTGTGAGTGGGCGTGAATTCCTATGAGATGATTTATGGAGA 918
Db 841 GGCCTGAGCTGTGACTGTGGTGTGAGTGGGCGTGAATTCCTATGAGATGATTTATGGAGA 900
QY 919 TCCCTCTCCGAGAGGAACTCTGCCAGAACTTCAATAACATTAATGAATTTCCAGCGG 978
Db 901 TCCCTCTCCGAGAGGAACTCTGCCAGAACTTCAATAACATTAATGAATTTCCAGCGG 960
QY 979 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1038
Db 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020
QY 1039 TTGTTGCGCGGCAAGAAAGAGAGAGCTGAAGTTTGAAGGTCTTTGCTGCCATCTCTTTCTTC 1098
Db 1021 TTGTTGCGCGGCAAGAAAGAGAGAGCTGAAGTTTGAAGGTCTTTGCTGCCATCTCTTTCTTC 1080
QY 1099 TCTAAAATTTGACTGGAACAACTTCGTAACCTCTCCCTCCCTTCCTGTTTCCACCCCTCAAG 1158
Db 1081 TCTAAAATTTGACTGGAACAACTTCGTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
QY 1159 TCTGACGATGACACCTCCAAATTTTGATGAACCAAGAGAAAGTTCGTTGGGTTTCATCTCT 1218
Db 1141 TCTGACGATGACACCTCCAAATTTTGATGAACCAAGAGAAAGTTCGTTGGGTTTCATCTCTCT 1200
QY 1219 CCGTCCAGCTGAGCCCTCAGCTTCTCGGGTGAAGAACTGCGGTTTGGGGTTTTTCG 1278
Db 1201 CCGTCCAGCTGAGCCCTCAGCTTCTCGGGTGAAGAACTGCGGTTTGGGGTTTTTCG 1260
QY 1279 TACAGCAAGCACTGGGGATTCTTGGTAGATCTGAGTCTGTTGTCGGGTCTGGACTCC 1338
```

Db 1261 TACAGCAAGGCACTGGGGAATCTTGGTAGATCTGAGTCTGTGTCGGGTCTGGACTCC 1320  
Qy 1339 CTTGCCAAGACTAGCTCCATGGAAAGAACTTCTCATCAAAAAGCAAGAGCTACAGAC 1398  
Db 1321 CTTGCCAAGACTAGCTCCATGGAAAGAACTTCTCATCAAAAAGCAAGAGCTACAGAC 1380  
Qy 1399 TCTCAGACAAAGTGTACACAGATGGACAGGAATGACCCGGTATCATCGGAGGTGCA 1458  
Db 1381 TCTCAGACAAAGTGTACACAGATGGACAGGAATGACCCGGTATCATCGGAGGTGCA 1440  
Qy 1459 GAGTGGAGGCTGTGCTTAGTCAGAAAGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1518  
Db 1441 GAGTGGAGGCTGTGCTTAGTCAGAAAGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1500  
Qy 1519 TCCCTCTCGAGCAGGACCTTGCTACTATCATACAGAAATGAGTAGCTTAAGCGAAGT 1578  
Db 1501 TCCCTCTCGAGCAGGACCTTGCTACTATCATACAGAAATGAGTAGCTTAAGCGAAGT 1560  
Qy 1579 TTGAGCAAGCAGGATGGAGGTGTCCAGAGGATGACAAAGCACTGCAGCTTCTCCAT 1638  
Db 1561 TTGAGCAAGCAGGATGGAGGTGTCCAGAGGATGACAAAGCACTGCAGCTTCTCCAT 1620  
Qy 1639 GATATCAGAGCAGGACCGGAAGCTCCAAAGAAATCAAAGAGCAGGAGTACAGGCTCAA 1698  
Db 1621 GATATCAGAGCAGGACCGGAAGCTCCAAAGAAATCAAAGAGCAGGAGTACAGGCTCAA 1680  
Qy 1699 GTGGAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1758  
Db 1681 GTGGAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740  
Qy 1759 CGGAGTGATCTTACGAATCTGAGCTGAGAGACTCTCGGCTGCTCTCAAGAAATTCAG 1818  
Db 1741 CGGAGTGATCTTACGAATCTGAGCTGAGAGACTCTCGGCTGCTCTCAAGAAATTCAG 1800  
Qy 1819 CGGAACCGACAGATCTCAGCAATAACTTGTGAAGCTTAAGATCAAGGGAGCTGGA 1878  
Db 1801 CGGAACCGACAGATCTCAGCAATAACTTGTGAAGCTTAAGATCAAGGGAGCTGGA 1860  
Qy 1879 GTGGGGAATATCGAACTGGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938  
Db 1861 GTGGGGAATATCGAACTGGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
Qy 1939 CTCAGAGAAACTGGAGAAGCTG----- 1963  
Db 1921 CTCAGAGAAACTGGAGAAGCTGTAAGCCGACGAGGCCACCGAGCTGTGTCAG 1980  
Qy 1964 -----CAAAGGAGCAGCGAGAGGAGCTGGAGAAGCTGAGAACTGAGAAACCGAGAG 2010  
Db 1981 AATATCCGCGAGCAAGAGCCGAGCGAGGAGCTGGAGAGCTGAGAACTGAGAAACCGAGAG 2040  
Qy 2011 GATTTCTTGAAGGCATCAGAAAGAGCTGTGGAAAGCTGAGAAACCGCGCCATCTCTG 2070  
Db 2041 GATTTCTTGAAGGCATCAGAAAGAGCTGTGGAAAGCTGAGAAACCGCGCCATCTCTG 2100  
Qy 2071 GAGAAACAAGTAAAGAGACTAGAGACCATGGACCTAGAGAAACAGACTGAAGATGAC 2130  
Db 2101 GAGAAACAAGTAAAGAGACTAGAGACCATGGACCTAGAGAAACAGACTGAAGATGAC 2160  
Qy 2131 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATATAATTTCTGGAGCTCGAAGAG 2190  
Db 2161 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATATAATTTCTGGAGCTCGAAGAG 2220  
Qy 2191 AATCATCGGAGCCCAAGCTCTCAGCCAGCAGCTAGAGTGCACCTGTAACAGAAAGAG 2250  
Db 2221 AATCATCGGAGCCCAAGCTCTCAGCCAGCAGCTAGAGTGCACCTGTAACAGAAAGAG 2280  
Qy 2251 CAGCACTATGAGGAAAGATTAAGTGTGGCAATTCAGATAAAGAAAGACCTGGCTGAC 2310  
Db 2281 CAGCACTATGAGGAAAGATTAAGTGTGGCAATTCAGATAAAGAAAGACCTGGCTGAC 2340  
Qy 2311 AAGGACACTGAGAACATGTCAGAGACACAGGAGGAGGCCCTATGAGAGGCCAA 2370

Db 2341 AAGGACACTGAGAAACATGATCGAGAGACACAGGAGGAGGCCCATGAGAGGCCAA 2400  
Qy 2371 ATTCTCAGCGAACAGAGGCGGATGATCAATGCTATGATTTCCAAAGATCAGATCCCTGGAA 2430  
Db 2401 ATTCTCAGCGAACAGAGGCGGATGATCAATGCTATGATTTCCAAAGATCAGATCCCTGGAA 2460  
Qy 2431 CAGAGGATTGTGGAACCTGTCTGAAAGCAATAAACTTCAGCAAAATAGCAGTCTTTTACC 2490  
Db 2461 CAGAGGATTGTGGAACCTGTCTGAAAGCAATAAACTTCAGCAAAATAGCAGTCTTTTACC 2520  
Qy 2491 CAAAGCAATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC 2550  
Db 2521 CAAAGCAATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC 2580  
Qy 2551 CTGAGACACAGGCTGGGAGTTGGAGGCCAGAACCGAANAATGGAGGAGCAGCTGGAG 2610  
Db 2581 CTGAGACACAGGCTGGGAGTTGGAGGCCAGAACCGAANAATGGAGGAGCAGCTGGAG 2640  
Qy 2611 AAGATCAAGCACCAGACCCAGTGAACAAGATCGGCTGCTGGAACTTGAGACAAGATTG 2670  
Db 2641 AAGATCAAGCACCAGACCCAGTGAACAAGATCGGCTGCTGGAACTTGAGACAAGATTG 2700  
Qy 2671 CGGAGGTCAGTCTAGAGCAGGAGCAGAAACTGAGAGCTCAAGCGCCAGCTCACAGAG 2730  
Db 2701 CGGAGGTCAGTCTAGAGCAGGAGCAGAAACTGAGAGCTCAAGCGCCAGCTCACAGAG 2760  
Qy 2731 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCCCTGAGGCTGCACGCGCG 2790  
Db 2761 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCCCTGAGGCTGCACGCGCG 2820  
Qy 2791 GCCCTGAGAGCAGCTTCCGAGCGAAGACAGAGCTGGAAGAGACCAAGAGAGCT 2850  
Db 2821 GCCCTGAGAGCAGCTTCCGAGCGAAGACAGAGCTGGAAGAGACCAAGAGAGCT 2880  
Qy 2851 GAAGAGAGATCCAGGCACTCAGCGCATAGAGATGAAATCAGCGCAAAATTTGATGCT 2910  
Db 2881 GAAGAGAGATCCAGGCACTCAGCGCATAGAGATGAAATCAGCGCAAAATTTGATGCT 2940  
Qy 2911 CTTCTAACAGCTGTAATCAAGACTGAGAGGAGCAGCTAAACAGAGTGCAGCGAG 2970  
Db 2941 CTTCTAACAGCTGTAATCAAGACTGAGAGGAGCAGCTAAACAGAGTGCAGCGAG 3000  
Qy 2971 GACAACTGAGCTCAACAACTTCTACTTGTCCAAACAACTCCGATGAGGCTTCT 3030  
Db 3001 GACAACTGAGCTCAACAACTTCTACTTGTCCAAACAACTCCGATGAGGCTTCT 3060  
Qy 3031 GGCGCCAACGAGATTTGACAACTCGAAGTGAAGTGAGCCATCTCCGCGGAGATC 3090  
Db 3061 GGCGCCAACGAGATTTGACAACTCGAAGTGAAGTGAGCCATCTCCGCGGAGATC 3120  
Qy 3091 ACGGAAACGAGAGTGCAGCTTACAGCAGAGCAAGCAAGATGAGGCTCTGAGACCAAG 3150  
Db 3121 ACGGAAACGAGAGTGCAGCTTACAGCAGAGCAAGCAAGATGAGGCTCTGAGACCAAG 3180  
Qy 3151 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA 3210  
Db 3181 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA 3240  
Qy 3211 AAAGAGCGGAGTGGGAGGCTTGGAGGAGCTCTCGGCTGATGAGAAATCCGAGTTTGGAG 3270  
Db 3241 AAAGAGCGGAGTGGGAGGCTTGGAGGAGCTCTCGGCTGATGAGAAATCCGAGTTTGGAG 3300  
Qy 3271 TGTGGGTTTGAGAGCTGAGAGATCTGCGACACCGAGAAACAGAGCAGCGCGAGAGCC 3330  
Db 3301 TGTGGGTTTGAGAGCTGAGAGATCTGCGACACCGAGAAACAGAGCAGCGCGAGAGCC 3360  
Qy 3331 GATCAGCGGATCACCGAGTCTCGCGAGGCTGGAGCTGCGAGTGAAGAGCAGAGGCT 3390  
Db 3361 GATCAGCGGATCACCGAGTCTCGCGAGGCTGGAGCTGGAGTGAAGAGCAGAGGCT 3420  
Qy 3391 GAGATTCTCCCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCT 3450  
Db 3421 GAGATTCTCCCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCT 3480

QY 3451 GACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGATGCCGGAAGCTTA 3510  
Db 3481 GACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGATGCCGGAAGCTTA 3540  
QY 3511 CAGCAGAGCTGAGAGCTGAAGAGAGCTCAACAGAGCTTCTGGAAGAGCAAGCCAAA 3570  
Db 3541 CAGCAGAGCTGAGAGCTGAAGAGAGCTCAACAGAGCTTCTGGAAGAGCAAGCCAAA 3600  
QY 3571 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTCCGTCCTGACTCAAGGACTGCAA 3630  
Db 3601 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTCCGTCCTGACTCAAGGACTGCAA 3660  
QY 3631 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGAGTATCAGCTG 3690  
Db 3661 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGAGTATCAGCTG 3720  
QY 3691 GAAAAATTCAGGTTCTCTATTTCTCATGAAAAAGTGAATGAAAGGCACTATTTCTCAA 3750  
Db 3721 GAAAAATTCAGGTTCTCTATTTCTCATGAAAAAGTGAATGAAAGGCACTATTTCTCAA 3780  
QY 3751 CAAACCAAATCATTTGATTTTTCGAAGCCAAATGGACCAACCTGCTTAAAGAAAAAG 3810  
Db 3781 CAAACCAAATCATTTGATTTTTCGAAGCCAAATGGACCAACCTGCTTAAAGAAAAAG - - 3838  
QY 3811 GGTTTATTTAGTCAGGAAAGAGGACCTGCTTTTACCCACACAGGTTTCTCTGAGTAC 3870  
Db 3839 -----AGTTTCTCTGAGTAC 3855  
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTAGAGAAAGCC 3930  
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTAGAGAAAGCC 3915  
QY 3931 CTTTCAGAGACCGCATCGAGCTCGGTCGCGCCCGGAGGAGAGCTGCCACCGCAAGCA 3990  
Db 3916 CTTTCAGAGACCGCATCGAGCTCGGTCGCGCCCGGAGGAGAGCTGCCACCGCAAGCA 3975  
QY 3991 ACGGACCAACCAACCCATCCACGCCAGCCACCGGAGGACAGCATGCCATGTCGCC 4050  
Db 3976 ACGGACCAACCAACCCATCCACGCCAGCCACCGGAGGACAGCATGCCATGTCGCC 4035  
QY 4051 ATCGTCGGTCCAGAGACACAGCCGAGTGCATGAGCTGCTGGCCCGGCATTCAGC 4110  
Db 4036 ATCGTCGGTCCAGAGACACAGCCGAGTGCATGAGCTGCTGGCCCGGCATTCAGC 4095  
QY 4111 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAAGGAAAGCATGCAC 4170  
Db 4096 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAAGGAAAGCATGCAC 4155  
QY 4171 CACAATATTCCTCAACGATTCACGTAAGACTGAACATGCGAGGCCAAGAGTGTGCTGTG 4230  
Db 4156 CACAATATTCCTCAACGATTCACGTAAGACTGAACATGCGAGGCCAAGAGTGTGCTGTG 4215  
QY 4231 TGCTCTGGATACCGTGCACCTTTGGAGCCAGGCATCCAAATGCTCGAATGTCAGGTGATG 4290  
Db 4216 TGCTCTGGATACCGTGCACCTTTGGAGCCAGGCATCCAAATGCTCGAATGTCAGGTGATG 4275  
QY 4291 TGTCAACCCCAAGTGTCTCAACGTCGTGTCAGCCACCTCGCGGTTGCCCTGCTGAATATGCC 4350  
Db 4276 TGTCAACCCCAAGTGTCTCAACGTCGTGTCAGCCACCTCGCGGTTGCCCTGCTGAATATGCC 4335  
QY 4351 ACACATTCACCGAGCCCTTCTGCGGTGACAAATGAATGCCAGGTCCTCAGACCAAG 4410  
Db 4336 ACACATTCACCGAGCCCTTCTGCGGTGACAAATGAATGAATGCCAGGTCCTCAGACCAAG 4395  
QY 4411 GAGCCACGAGCAGCTTGCACTCGAAGGAGGTGGAATGAAAGGTGCCAGGAATAACAAACGA 4470  
Db 4396 GAGCCACGAGCAGCTTGCACTCGAAGGAGGTGGAATGAAAGGTGCCAGGAATAACAAACGA 4455  
QY 4471 GGACAGCAAGGCTGGGACAGGAAGTACATTTGCTGGAGGAGTCAAAAGTCTCATTTAT 4530  
Db 4456 GGACAGCAAGGCTGGGACAGGAAGTACATTTGCTGGAGGAGTCAAAAGTCTCATTTAT 4515

QY 4531 GACAATGAAGCCAGAGAGAGCTGGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTTTCCC 4590  
Db 4516 GACAATGAAGCCAGAGAGAGCTGGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTTTCCC 4575  
QY 4591 GACGGGATGTATCTATTTTCATGTGTGCGTTGGTCTTCCGAATCAGCAATCAGCCCCAAA 4650  
Db 4576 GACGGGATGTATCTATTTTCATGTGTGCGTTGGTCTTCCGAATCAGCAATCAGCCCCAAA 4635  
QY 4651 GCA----- 4653  
Db 4636 GCAGATGTCCCATACATACCTACTGAAGATGGAATCTCACCCGCACACCACTGCTGTCGCCCGG 4695  
QY 4654 ----- 4653  
Db 4696 AGAACCTCTACTTGTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4755  
QY 4654 -----GAAAAAGCAGAAGCTGATGCTAAACTG 4680  
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATGCTAACTG 4815  
QY 4681 CTTGGAAAACTCCCTGCTGAAAACTGGAAGGTGATGACGCTTAGACATGAATGACGCTG 4740  
Db 4816 CTTGGAAAACTCCCTGCTGAAAACTGGAAGGTGATGACGCTTAGACATGAATGACGCTG 4875  
QY 4741 CCCTTCAGTCACAGGTGTGTGTTGGGCAACGAGGAAGGGCTCTACGCCCTGAATGTC 4800  
Db 4876 CCCTTCAGTCACAGGTGTGTGTTGGGCAACGAGGAAGGGCTCTACGCCCTGAATGTC 4935  
QY 4801 TTGAAAACTCCCTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860  
Db 4936 TTGAAAACTCCCTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4995  
QY 4861 AAGGACCTGAGAGAGCTACTCATGATAGCAGAGAGAGGGGCACTGTCTCTGTGGAC 4920  
Db 4996 AAGGACCTGAGAGAGCTACTCATGATAGCAGAGAGAGGGGCACTGTCTCTGTGGAC 5055  
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4980  
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 5115  
QY 4981 CCCAACATTTTGAAGCTGTCAAGGGTGCCTTGTGTTGGGCGAGCAAGATTGAGAAC 5040  
Db 5116 CCCAACATTTTGAAGCTGTCAAGGGTGCCTTGTGTTGGGCGAGCAAGATTGAGAAC 5175  
QY 5041 GGGCTCTGCATCTGTGCAGCATGCCAGCAAAAGTGTCTATTCTCGCTTACAAACGAAAAAC 5100  
Db 5176 GGGCTCTGCATCTGTGCAGCATGCCAGCAAAAGTGTCTATTCTCGCTTACAAACGAAAAAC 5235  
QY 5101 CTCAGCAATATCTGCATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5160  
Db 5236 CTCAGCAATATCTGCATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5295  
QY 5161 TTCACCAATTACAGTATCCCTCATTTGGAAACCAATAAATTTCTACGAATCGACATGAAGCAG 5220  
Db 5296 TTCACCAATTACAGTATCCCTCATTTGGAAACCAATAAATTTCTACGAATCGACATGAAGCAG 5355  
QY 5221 TACACGCTCGAGGAATTCCTGGATAAGAATGACCATTCCTTGGCAGCTGTGTGTTGCC 5280  
Db 5356 TACACGCTCGAGGAATTCCTGGATAAGAATGACCATTCCTTGGCAGCTGTGTGTTGCC 5415  
QY 5281 GCCTCTTCCAGAGCTTCCCTGTCTCAATCGTCAGAGTGAACAGCCGAGGGCAGCAGAG 5340  
Db 5416 GCCTCTTCCAGAGCTTCCCTGTCTCAATCGTCAGAGTGAACAGCCGAGGGCAGCAGAG 5475  
QY 5341 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAACCGTAGC 5400  
Db 5476 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAACCGTAGC 5535  
QY 5401 CGCAGACAGATCTCAAGTGGAGTCCGCTTACCTTTGGCTTTGCTTACAGAGAACCTTAT 5460  
Db 5536 CGCAGACAGATCTCAAGTGGAGTCCGCTTACCTTTGGCTTTGCTTACAGAGAACCTTAT 5595  
QY 5461 CTGTTTGTGACCCACTTCACTCACTCGAAGTAATTTGAGATCCAGGCAACGCTCTCTCAGCA 5520

```
Db 5596 CTTGTTGTGACCCACCTTCAACTCACTCGAAGTAATTTGAGATCGAGCAGCTTCTCTCAGCA 5655
|
|
|
Qy 5521 GGGACCCCTGCCGAGCGTACTTGGAACATCCCGAACCCCGCTACTGCGGCCCTGCCATT 5580
|
|
|
Db 5656 GGGACCCCTGCCGAGCGTACTTGGAACATCCCGAACCCCGCTACTGCGGCCCTGCCATT 5715
|
|
|
Qy 5581 TCTCTAGGAGCGATTTACCTTGGGCTCCTCATACCGAGTAATAATTAAGGTCATTTGCTGC 5640
|
|
|
Db 5716 TCTCTAGGAGCGATTTACCTTGGGCTCCTCATACCGAGTAATAATTAAGGTCATTTGCTGC 5775
|
|
|
Qy 5641 AAGGGAAACCTCGTGAAGAGTCCGGCACTGAAACACCCAGCGGGCCCGCTCCACCTCCCGC 5700
|
|
|
Db 5776 AAGGGAAACCTCGTGAAGAGTCCGGCACTGAAACACCCAGCGGGCCCGCTCCACCTCCCGC 5835
|
|
|
Qy 5701 AGAGGCCCAAGAGCGAGGCCACCCACGTCATCAACGAGCAGATACCAAGCGGTGGCC 5760
|
|
|
Db 5836 AGAGGCCCAAGAGCGAGGCCACCCACGTCATCAACGAGCAGATACCAAGCGGTGGCC 5895
|
|
|
Qy 5761 TCCAGGCCAGCGCGCCCGAAGGCCCGAGCCACCCCGGAGAGCCAGCACACCCACCGC 5820
|
|
|
Db 5896 TCCAGGCCAGCGCGCCCGAAGGCCCGAGCCACCCCGGAGAGCCAGCACACCCACCGC 5955
|
|
|
Qy 5821 TACCGGAGGGCGGACCGAGCTGCGAGGAGCAAGTCTCTGCGCCCGCCCTCGAGGCGA 5880
|
|
|
Db 5956 TACCGGAGGGCGGACCGAGCTGCGAGGAGCAAGTCTCTGCGCCCGCCCTCGAGGCGA 6015
|
|
|
Qy 5881 GAGAGTCCCGCGCGGAGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5940
|
|
|
Db 6016 GAGAGTCCCGCGCGGAGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 6075
|
|
|
Qy 5941 GAAGACAGCAGCAGGGCGCGCTCGCTGCGGAGCGCTGAGGACCCCGCTGCCAGGTG 6000
|
|
|
Db 6076 GAAGACAGCAGCAGGGCGCGCTCGCTGCGGAGCGCTGAGGACCCCGCTGCCAGGTG 6135
|
|
|
Qy 6001 AACAGGGAGAGGGCAGAGTGCCTCTCAAGTTTTCACGGTTTAAACACTGCTCACTATTAT 6060
|
|
|
Db 6136 AACAGGTCTGGACCACTCTTCAGTATAAATCTCAGCCAGGAAACCAACTCTCATCT 6195
|
|
|
RESULT 5
LOCUS AX671112 6156 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 8 from Patent WO03004523.
ACCESSION AX671112
VERSION AX671112.1 GI:29329572
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Zhu, Z.
REFERENCE
AUTHORS Regulation of human citron rho/rac-interacting kinase
TITLE Patent: WO 03004523-A 8 16-JAN-2003;
JOURNAL Bayer Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
source
1..6156
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 86.28; Score 5666.6; DB 6; Length 6156;
Best Local Similarity 95.94; Pred. No. 0;
Matches 5945; Conservative 0; Mismatches 9; Indels 243; Gaps 3;
Qy 19 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGGTCTGCTGAACCCATT 78
|
|
|
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGGTCTGCTGAACCCATT 60
|
|
|
Qy 79 GCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATGACTCAA 138
|
|
|
Db 1141 TCTGACGATGACACCTCCAAATTTTGTATGAACACGAGAGAAATTCGTGGGTTTCATCCTCT 1200
|
|
|
```

```
Db 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCCCTTTATGACTCAA 120
|
|
|
Qy 139 CAGCAGATGCTCTCTCTTTCCCGAGAGGATATTAGATCCCTCTTTGTTCTCTTTGAA 198
|
|
|
Db 121 CAGCAGATGCTCTCTCTTTCCCGAGAGGATATTAGATCCCTCTTTGTTCTCTTTGAA 180
|
|
|
Qy 199 GAATSCAGTCAGCGCTGCTCTGATGAAGATTAAAGCAGTCGAGCAACTTTGTCGGGAAGTAT 258
|
|
|
Db 181 GAATSCAGTCAGCGCTGCTCTGATGAAGATTAAAGCAGTCGAGCAACTTTGTCGGGAAGTAT 240
|
|
|
Qy 259 TCCGACACCATAGCTAGTTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 318
|
|
|
Db 241 TCCGACACCATAGCTAGTTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 300
|
|
|
Qy 319 AGTCTTCTAGCTTGTGGTCACTTTTGTCTGAAGTCAGGTGTTAAGAGAGAAAGCAACCGGG 378
|
|
|
Db 301 AGTCTTCTAGCTTGTGGTCACTTTTGTCTGAAGTCAGGTGTTAAGAGAGAAAGCAACCGGG 360
|
|
|
Qy 379 GACATCTATGCTATGAAGTGTAGAGAGAGAGCTTTTATGCGCCAGAGAGAGGTTTCA 438
|
|
|
Db 361 GACATCTATGCTATGAAGTGTAGAGAGAGAGCTTTTATGCGCCAGAGAGAGGTTTCA 420
|
|
|
Qy 439 TTTTTCGAGGAAGCGGAAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 498
|
|
|
Db 421 TTTTTCGAGGAAGCGGAAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480
|
|
|
Qy 499 CAGTATCCCTTTCAGGACAAAAATCACTTTTATCTGATGAGGAAATATCAAGCTTGAGGG 558
|
|
|
Db 481 CAGTATCCCTTTCAGGACAAAAATCACTTTTATCTGATGAGGAAATATCAAGCTTGAGGG 540
|
|
|
Qy 559 GACTTGTGTCACCTTTTGAATAGATAGGACCACTTATGATGAAACCTGATACAGTTT 618
|
|
|
Db 541 GACTTGTGTCACCTTTTGAATAGATAGGACCACTTATGATGAAACCTGATACAGTTT 600
|
|
|
Qy 619 TACCTAGCTAGCTGATTTTGGCTGTTTCAAGCGTTTATCTGATGGGATACGTGCATCGA 678
|
|
|
Db 601 TACCTAGCTAGCTGATTTTGGCTGTTTCAAGCGTTTATCTGATGGGATACGTGCATCGA 660
|
|
|
Qy 679 GACATCAAGCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 738
|
|
|
Db 661 GACATCAAGCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
|
|
|
Qy 739 GGATCTCGCGGAAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGGATTTGGGACC 798
|
|
|
Db 721 GGATCTCGCGGAAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGGATTTGGGACC 780
|
|
|
Qy 799 CCAGATTACATGGCTCCTGAAAGTGTGACTGTGATGAACCGGGATGGAAGAGCACCTAC 858
|
|
|
Db 781 CCAGATTACATGGCTCCTGAAAGTGTGACTGTGATGAACCGGGATGGAAGAGCACCTAC 840
|
|
|
Qy 859 GGCTTGAGCTGTGACTGCTGCTGAGTGGGCTGATTGCTCTATCAGATGATTTATGGGAGA 918
|
|
|
Db 841 GGCTTGAGCTGTGACTGCTGCTGAGTGGGCTGATTGCTCTATCAGATGATTTATGGGAGA 900
|
|
|
Qy 919 TCCCGCTTCGAGAGGAAACCTCTGCGAGAACCTTCAATTAACATTATGAATTTCCAGCGG 978
|
|
|
Db 901 TCCCGCTTCGAGAGGAAACCTCTGCGAGAACCTTCAATTAACATTATGAATTTCCAGCGG 960
|
|
|
Qy 979 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGTGACTGCTTTCTTGTATCTGATTTCAAGC 1038
|
|
|
Db 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGTGACTGCTTTCTTGTATCTGATTTCAAGC 1020
|
|
|
Qy 1039 TTGTTGTGCGGCCAGAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATTCCTTCTTC 1098
|
|
|
Db 1021 TTGTTGTGCGGCCAGAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATTCCTTCTTC 1080
|
|
|
Qy 1099 TCTAAAATTGACTGGAACAACATTCGTAATCTCTCCCTCCCTCGTCCACCTCAAG 1158
|
|
|
Db 1081 TCTAAAATTGACTGGAACAACATTCGTAATCTCTCCCTCCCTCGTCCACCTCAAG 1140
|
|
|
Qy 1159 TCTGACGATGACACTCCAATTTTGTATGAACACGAGAGAAATTCGTGGGTTTCACTCT 1218
|
|
|
Db 1141 TCTGACGATGACACTCCAATTTTGTATGAACACGAGAGAAATTCGTGGGTTTCACTCTCT 1200
|
|
|
```

QY 1219 CCGTGCCAGCTAGAGCCCTCAGGCTTCTCGGTGAAGAACTGCCGTTTGTGGGTTTTCG 1278  
Db 1201 CCGTGCCAGCTAGAGCCCTCAGGCTTCTCGGTGAAGAACTGCCGTTTGTGGGTTTTCG 1260  
QY 1279 TACAGCAAGGCACTGGGATTTCTTGCTAGATCTGAGTCTGTGTTGTCGGGTCTGAGCTCC 1338  
Db 1261 TACAGCAAGGCACTGGGATTTCTTGCTAGATCTGAGTCTGTGTTGTCGGGTCTGAGCTCC 1320  
QY 1339 CTTGCAAGACTAGCTCAGTGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1398  
Db 1321 CTTGCAAGACTAGCTCAGTGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1380  
QY 1399 TCTCAGGCAAGAGTGCAGAAAGTGCAGAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1458  
Db 1381 TCTCAGGCAAGAGTGCAGAAAGTGCAGAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1440  
QY 1459 GAGGTGAGGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518  
Db 1441 GAGGTGAGGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
QY 1519 TCCCTCCTGGAGCAGGACTTGTACTACTACATCAAGAAATGCAGATGCAGTAGCTTAAAGCGAGT 1578  
Db 1501 TCCCTCCTGGAGCAGGACTTGTACTACTACATCAAGAAATGCAGATGCAGTAGCTTAAAGCGAAGT 1560  
QY 1579 TTGGAGCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTCAGCTTCTCCAT 1638  
Db 1561 TTGGAGCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTCAGCTTCTCCAT 1620  
QY 1639 GATATCAGAGCAGAGCGGAGCTCCAGAAATCAAGAGCAGGAGTACAGGCTCAA 1698  
Db 1621 GATATCAGAGCAGAGCGGAGCTCCAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680  
QY 1699 GTGGAAGAAATGAGTGTGATGATGAATCAGTTGGAAGAGGATCTTGCTCAGCAAGAGA 1758  
Db 1681 GTGGAAGAAATGAGTGTGATGATGAATCAGTTGGAAGAGGATCTTGCTCAGCAAGAGA 1740  
QY 1759 CGAGTGATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1818  
Db 1741 CGAGTGATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1800  
QY 1819 CGGAAGCGACAGATGTCAGATAAATGTTGAAGGCTTAGGATCAAGGAGACCTGAA 1878  
Db 1801 CGGAAGCGACAGATGTCAGATAAATGTTGAAGGCTTAGGATCAAGGAGACCTGAA 1860  
QY 1879 GTGGAGAAATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1938  
Db 1861 GTGGAGAAATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920  
QY 1939 CTCGAAGGAACTGGAGAGCT----- 1962  
Db 1921 CTCGAAGGAACTGGAGAGCTGTAAAGCCAGCACGGAGGCCACCGAGCTGTCGAG 1980  
QY 1963 -----GCAAGGAGCGAGCCGAGAGGAGCTGGAGAACTGCAGAACCGAGAG 2010  
Db 1981 AATATCCCGCAGGCAAGAGCGAGCCGAGAGGAGCTGGAGAGCTGCAGAACCGAGAG 2040  
QY 2011 GATTTCTGAGGATCAGAAAGAGCTGTGGAGCTGAGGAACCGCCGCTTCTCTG 2070  
Db 2041 GATTTCTGAGGATCAGAAAGAGCTGTGGAGCTGAGGAACCGCCGCTTCTCTG 2100  
QY 2071 GAGAACAGGTAAAGAGCTAGAGACCATGAGCCGTAGAGAAACAGAGTGAAGATGAC 2130  
Db 2101 GAGAACAGGTAAAGAGCTAGAGACCATGAGCCGTAGAGAAACAGAGTGAAGATGAC 2160  
QY 2131 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGGAGCTGAAGAG 2190  
Db 2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGGAGCTGAAGAG 2220  
QY 2191 AACAATCGGAGGCCAGTCTCAGCCAGCAGCTAGAGTGCACCTGAACAGAGAGAG 2250  
Db 2221 AACAATCGGAGGCCAGTCTCAGCCAGCAGCTAGAGTGCACCTGAACAGAGAGAG 2280

2251 CAGCACTATCAGAAAGATTAAAGTGTGGACAAATCAGATTAAGAAAGACCTGGCTGAC 2310  
Db 2281 CAGCACTATCAGAAAGATTAAAGTGTGGACAAATCAGATTAAGAAAGACCTGGCTGAC 2340  
QY 2311 AAGGAGACACTGGAGAACATGATGCAGAGACACGAGAGAGAGGCCCATGATGAAGGGCAAA 2370  
Db 2341 AAGGAGACACTGGAGAACATGATGCAGAGACACGAGAGAGAGGCCCATGATGAAGGGCAAA 2400  
QY 2371 ATTCTCAGCGAAACAGAAAGGCGATGATCAATGCTTATGGATTCAGAGTCAGATCCCTGGAA 2430  
Db 2401 ATTCTCAGCGAAACAGAAAGGCGATGATCAATGCTTATGGATTCAGAGTCAGATCCCTGGAA 2460  
QY 2431 CAGAGATTGTGAAACTGTCTGAAGCCAAATAAACTTTGCAGCAAAATAGCAGTCTTTTACC 2490  
Db 2461 CAGAGATTGTGAAACTGTCTGAAGCCAAATAAACTTTGCAGCAAAATAGCAGTCTTTTACC 2520  
QY 2491 CAAAGGAACATGAAGGCCCAAAGAGAGATGATTTCTGAATCTCAGGCAAACAGAAATTTTAC 2550  
Db 2521 CAAAGGAACATGAAGGCCCAAAGAGAGATGATTTCTGAATCTCAGGCAAACAGAAATTTTAC 2580  
QY 2551 CTGGAGACACAGCTGGGAAGTTGGAGGCCCAAGACCGAAACTTGGAGGAGCAGCTGGAG 2610  
Db 2581 CTGGAGACACAGCTGGGAAGTTGGAGGCCCAAGACCGAAACTTGGAGGAGCAGCTGGAG 2640  
QY 2611 AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGGCTGCTGAACTGGAGACAAGATTG 2670  
Db 2641 AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGGCTGCTGAACTGGAGACAAGATTG 2700  
QY 2671 CGGAGGTGAGTCTAGAGCAGGAGGACAGAACTGGAGCTCAAGCCAGCTCAGAG 2730  
Db 2701 CGGAGGTGAGTCTAGAGCAGGAGGACAGAACTGGAGCTCAAGCCAGCTCAGAG 2760  
QY 2731 CTACAGCTCTCCCTGCGAGGACGCGAGTCACTGTTGACAGCCCTGCGAGGCTGACCGGCG 2790  
Db 2761 CTACAGCTCTCCCTGCGAGGACGCGAGTCACTGTTGACAGCCCTGCGAGGCTGACCGGCG 2820  
QY 2791 GCCTTGAGAGCAGGCTTCCGAGGCGAGACAGAGCTTGAAGAGACCAAGAGAGCT 2850  
Db 2821 GCCTTGAGAGCAGGCTTCCGAGGCGAGACAGAGCTTGAAGAGACCAAGAGAGCT 2880  
QY 2851 GAAGAGAGATCAGGCACTCAGGACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2910  
Db 2881 GAAGAGAGATCAGGCACTCAGGACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2940  
QY 2911 CTTCTGAACAGCTGTACTGTAAATCAAGAGCTGGAGAGCAGCTAAACAGCTGACCGAG 2970  
Db 2941 CTTCTGAACAGCTGTACTGTAAATCAAGAGCTGGAGAGCAGCTAAACAGCTGACCGAG 3000  
QY 2971 GACAAAGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3030  
Db 3001 GACAAAGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060  
QY 3031 GCGGCCAACAGCAGATTTGTACAACTGCGAAGTGAAGTGGACCAATCTCCGCCGGGAGATC 3090  
Db 3061 GCGGCCAACAGCAGATTTGTACAACTGCGAAGTGAAGTGGACCAATCTCCGCCGGGAGATC 3120  
QY 3091 ACGGAACGAGAGATGAGCTTACAGCCAGAGCAACAGTGGAGGCTCTGAGAGCACG 3150  
Db 3121 ACGGAACGAGAGATGAGCTTACAGCCAGAGCAACAGTGGAGGCTCTGGAAGACGAG 3180  
QY 3151 TGCAACATGCTGAGGAAACAGAGTCTATGGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3210  
Db 3181 TGCAACATGCTGAGGAAACAGAGTCTATGGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3240  
QY 3211 AAGAGCGGAGTGGAGGCTGGAGAGGCTCTGGGTGATGAGAAATCCCAAGTTTGGAG 3270  
Db 3241 AAGAGCGGAGTGGAGGCTGGAGAGGCTCTGGGTGATGAGAAATCCCAAGTTTGGAG 3300  
QY 3271 TGTCCGGTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGAGCGGCGAGGCC 3330  
Db 3301 TGTCCGGTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGAGCGGCGAGGCC 3360  
QY 3331 GATCAGCGGATCACCGAGTCTCGCCAGGCTGCTGGAGCTGGCAGTGAAGGAGCAACAGGCT 3390

Db 3361 GATCAGCGGATACCCAGTCTCCAGGTTGGTGGAGCTGGCAGTGAAGGAGCACAAAGGCT 3420  
QY GAGATTCTCGCTCTGAGCAGGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCCTCTCT 3450  
Db 3421 GAGATTCTCGCTCTGAGCAGGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCCTCTCT 3480  
QY GACAAGCTTAATGACCTGAGAGAGAGCATGCTATGCTTGAATGAATGCCGAGAGCTTGA 3510  
Db 3481 GACAAGCTTAATGACCTGAGAGAGAGCATGCTATGCTTGAATGAATGCCGAGAGCTTGA 3540  
QY CAGCAGAAAGCTGAGAGCTGAACGAGAGCTCAACAGAGAGCTTCTGGAAGAGCAAGCAAA 3570  
Db 3541 CAGCAGAAAGCTGAGAGCTGAACGAGAGCTCAACAGAGAGCTTCTGGAAGAGCAAGCAAA 3600  
QY TTACAGCAGCAGATGAGCCTGAGAGAGAAATCAATTTTCGGTCTGACTCAAGAGCTGCAA 3630  
Db 3601 TTACAGCAGCAGATGAGCCTGAGAGAGAAATCAATTTTCGGTCTGACTCAAGAGCTGCAA 3660  
QY GAAGCTCTAGATCGGCTGATCTACTGAAGAGAGAGAGTGAATGGAGTATCAGCTG 3690  
Db 3661 GAAGCTCTAGATCGGCTGATCTACTGAAGAGAGAGAGTGAATGGAGTATCAGCTG 3720  
QY GAAAAATTCAGGTTCTTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3750  
Db 3721 GAAAAATTCAGGTTCTTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3780  
QY CAACCAAACTCATTGATTTCTGCAAGCCAAATGGAACCACTGCTAAAGABAAAG 3810  
Db 3781 CAACCAAACTCATTGATTTCTGCAAGCCAAATGGAACCACTGCTAAAGABAAAG 3838  
QY GGTTTATTTAGTCGAGGAAAGAGGACCTGCTTTACCCACACAGAGTTCTCTGCAAGTAC 3870  
Db 3839 -----AGTTCCTCTGCAAGTAC 3855  
QY AATGAGCTGAAGTGGCCCTGGAGAGAGAGAAAGTCTGCTGTCAGAGCTAGAGAGGCC 3930  
Db 3856 AATGAGCTGAAGTGGCCCTGGAGAGAGAGAAAGTCTGCTGTCAGAGCTAGAGAGGCC 3915  
QY CTTTCAGAGACCCGATCGAGCTCCGCTCGCCGCGGAGAGAGCTGCCACCCGAAAGCA 3990  
Db 3916 CTTTCAGAGACCCGATCGAGCTCCGCTCGCCGCGGAGAGAGCTGCCACCCGAAAGCA 3975  
QY ACAGACACCCACACCCATCCAGCCAGGCAACCGGAGGAGAGAGTCCCATGTCGCGCC 4050  
Db 3976 ACAGACACCCACACCCATCCAGCCAGGCAACCGGAGGAGAGAGTCCCATGTCGCGCC 4035  
QY ATCGTGGGTTCGAGAGACAGCCAGTGCATGAGCTGCTGGCCCGCCATCCAGC 4110  
Db 4036 ATCGTGGGTTCGAGAGACAGCCAGTGCATGAGCTGCTGGCCCGCCATCCAGC 4095  
QY CGCAGAAAGAGGCTTCAACTCCAGAGGAAATTTAGTCGCGCTTTAAGGAAACGCAATGCA 4155  
Db 4096 CGCAGAAAGAGGCTTCAACTCCAGAGGAAATTTAGTCGCGCTTTAAGGAAACGCAATGCA 4170  
QY CACAATATTTCTCACCGATTCAGCTAGAGCTGAACATGGAGGAGCAGAGTGTGCTGTG 4230  
Db 4156 CACAATATTTCTCACCGATTCAGCTAGAGCTGAACATGGAGGAGCAGAGTGTGCTGTG 4215  
QY TGTCTGGATACCGTGGACCTTTGAGCCAGGATCCAAATGCTCGAATGTGCTGAGTGTG 4290  
Db 4216 TGTCTGGATACCGTGGACCTTTGAGCCAGGATCCAAATGCTCGAATGTGCTGAGTGTG 4275  
QY TGTCAACCCAGTGTCTCAGCTGCTTGCAGCAGCAGCTGCGGCTGCTGCTGCTGCTGCTG 4350  
Db 4276 TGTCAACCCAGTGTCTCAGCTGCTTGCAGCAGCAGCTGCGGCTGCTGCTGCTGCTGCTG 4335  
QY ACACACTTACCCAGGCTTCTCCGCTGACAAATGAACTCCCGCAGGCTTCCAGACCAAG 4410  
Db 4336 ACACACTTACCCAGGCTTCTCCGCTGACAAATGAACTCCCGCAGGCTTCCAGACCAAG 4395  
QY GAGCCAGCAGCAGCTTGCACCTTGGAGGGTGGATGAAGGTCGCCAGGAAATCAACAAAGA 4470

Db 4396 GAGCCAGCAGCAGCTTGCACCTGGAAGGGTGGATGAAGTGGCCAGGAATACAAACGA 4455  
QY GGACAGCAAGCTCGGACAGGAAGTACATTTGCTGAGGGATCAAAAGTCTCAITTTAT 4530  
Db 4456 GGACAGCAAGCTCGGACAGGAAGTACATTTGCTGAGGGATCAAAAGTCTCAITTTAT 4515  
QY GACATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCC 4590  
Db 4516 GACATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCC 4575  
QY GACGGGATGATCTATTCTATGCTGCGTTGGTCTCCGAACTCGCAATACAGCAAA 4650  
Db 4576 GACGGGATGATCTATTCTATGCTGCGTTGGTCTCCGAACTCGCAATACAGCAAA 4635  
QY GCA----- 4653  
Db 4636 GCAGATGCCATACACTGGAATGGAATCTCACCCGCACACCACTCTGCTGCGCGGG 4695  
QY ----- 4654  
Db 4696 AGAACCTCTACTTTGCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCTTA 4755  
QY -----GAAAAAGCAGAGCTGATGCTAACTG 4680  
Db 4756 GAATCAGTTGTCAGAGTGGGAGAGTTCTAGGGAAGAAAGCAGAGCTGATGCTAACTG 4815  
QY CTTGGAAACTCCCTGCTGAAACTGGAAGAGTGAACCGCTAGACATGAACTGCAAGCTG 4740  
Db 4816 CTTGGAAACTCCCTGCTGAAACTGGAAGAGTGAACCGCTAGACATGAACTGCAAGCTG 4875  
QY CCTTTCAGTACCAAGTGTGTTGGTGGCAGCCAGAGAGGGTCTACGCGCTGAATGTC 4800  
Db 4876 CCTTTCAGTACCAAGTGTGTTGGTGGCAGCCAGAGAGGGTCTACGCGCTGAATGTC 4935  
QY TTGAAAACTCCCTTAACCCATGTCCTCAGGAAATTTGGAGCAGTCTTCCAAATTTATATC 4860  
Db 4936 TTGAAAACTCCCTTAACCCATGTCCTCAGGAAATTTGGAGCAGTCTTCCAAATTTATATC 4995  
QY AAGGACCTGGAGAGTACTCATGATAGCAGAGAGAGCGGCACTGTGCTGTGGAC 4920  
Db 4996 AAGGACCTGGAGAGTACTCATGATAGCAGAGAGAGCGGCACTGTGCTGTGGAC 5055  
QY GTGAGAGAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGCACTCA 4980  
Db 5056 GTGAGAGAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGCACTCA 5115  
QY CCCAACATTTTGAAGCTGTCAAGGGTGCACCTTGTGTTGGGCGAGCAAGATTGGAAC 5040  
Db 5116 CCCAACATTTTGAAGCTGTCAAGGGTGCACCTTGTGTTGGGCGAGCAAGATTGGAAC 5175  
QY GGGCTCTGCATCTGTGAGCCATGCCAGCAAGTGTGCTATCTCCGCTACAGCAAGAAC 5100  
Db 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAGTGTGCTATCTCCGCTACAGCAAGAAC 5235  
QY CTCAGCAAACTACTGATCCCGAAAGAGATAGAGACCTCAGAGCCCTGACGCTGATCCAC 5160  
Db 5236 CTCAGCAAACTACTGATCCCGAAAGAGATAGAGACCTCAGAGCCCTGACGCTGATCCAC 5295  
QY TTCAACCAATTTAGTATCTCTATTGGAAACCAATTAATTTACGAATTCGACATGAGCAG 5220  
Db 5296 TTCAACCAATTTAGTATCTCTATTGGAAACCAATTAATTTACGAATTCGACATGAGCAG 5355  
QY TACAGCTTCGAGGAAATTCCTGGATGAAGTGAACCTTCTGCACTGTGCTGTGCTGCTG 5280  
Db 5356 TACAGCTTCGAGGAAATTCCTGGATGAAGTGAACCTTCTGCACTGTGCTGTGCTGCTG 5415  
QY GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAAGTGAACAGCGGCGAGCGAGAG 5281  
Db 5416 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAAGTGAACAGCGGCGAGCGAGAG 5340  
QY GAGTACTTGTGTTTCCAGGAAATTTGGAGTGTGCTGTTGATTTCTTACGGAACAGCTAGC 5400  
Db 5476 GAGTACTTGTGTTTCCAGGAAATTTGGAGTGTGCTGTTGATTTCTTACGGAACAGCTAGC 5535



Db 1081 TC1AAAATTGACTGGAACAACATTCGTAATCTCTCTCCCTTCGTTCCACCTCAAG 1140  
Qy 1159 TCTGACGATGACACCTCCAAATTTTGTATGAACCGAGAGAAGATTCGTGGGTTTCATCCTCT 1218  
Db 1141 TCTGACGATGACACCTCCAAATTTTGTATGAACCGAGAGAAGATTCGTGGGTTTCATCCTCT 1200  
Qy 1219 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGCAAGAACTGCGGTTTGGGGTTTGG 1278  
Db 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGCAAGAACTGCGGTTTGGGGTTTGG 1260  
Qy 1279 TACAGCAAGCACTGGGATCTTTGTGATGATCTGAGTCTTTGTGCGGTCTGGACTCC 1338  
Db 1261 TACAGCAAGCACTGGGATCTTTGTGATGATCTGAGTCTTTGTGCGGTCTGGACTCC 1320  
Qy 1339 CCTGCCAGACTAGCTCCATGGGAAGAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1398  
Db 1321 CCTGCCAGACTAGCTCCATGGGAAGAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1380  
Qy 1399 TCTCAGGACAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTACATCGGAGGTGTCA 1458  
Db 1381 TCTCAGGACAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTACATCGGAGGTGTCA 1440  
Qy 1459 GAGGTGAGGCTGTCTTAGTCAGAGGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1518  
Db 1441 GAGGTGAGGCTGTCTTAGTCAGAGGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500  
Qy 1519 TCCCTCTCGAGCAGGACCTTGCTACTCATCACAGAAATGCACTAGTCTTAAAGCGAAGT 1578  
Db 1501 TCCCTCTCGAGCAGGACCTTGCTACTCATCACAGAAATGCACTAGTCTTAAAGCGAAGT 1560  
Qy 1579 TTGAGCAAGCAAGATGGAGGTGTCTCCAGGAGATGCAAAAGCACTGAGCTTCTCCAT 1638  
Db 1561 TTGAGCAAGCAAGATGGAGGTGTCTCCAGGAGATGCAAAAGCACTGAGCTTCTCCAT 1620  
Qy 1639 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1698  
Db 1621 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1680  
Qy 1699 GTGGAAGAAATGAGGTGATGATGAATCAATTTGGAAGAGATCTTGTCTCAGCAAGAGA 1758  
Db 1681 GTGGAAGAAATGAGGTGATGATGAATCAATTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740  
Qy 1759 CCGAGTGAATCTTACGAATCTGAGCTGAGAGATCTCGGCTTGTCTGCTGAAGAAATTCAG 1818  
Db 1741 CCGAGTGAATCTTACGAATCTGAGCTGAGAGATCTCGGCTTGTCTGCTGAAGAAATTCAG 1800  
Qy 1819 CGGAAAGCGACAGAAATGTCAGCAATAAATTTGGAAGGCTAAGGATCAAGGGAAGCTGAA 1878  
Db 1801 CGGAAAGCGACAGAAATGTCAGCAATAAATTTGGAAGGCTAAGGATCAAGGGAAGCTGAA 1860  
Qy 1879 GTGGAGAAATATCGGAAATGGAAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1938  
Db 1861 GTGGAGAAATATCGGAAATGGAAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920  
Qy 1939 CTCGAGAGAAATCTGAGAGGCT----- 1962  
Db 1921 CTCGAGAGAAATCTGAGAGGCTGTAAAGCCAGCAGCGAGGCCACCGAGCTGCTGCAG 1980  
Qy 1963 -----CGAAGGAGCGAGCCGAGAGGAGCTGGAGAAAGCTGCAGAACCGAGAG 2010  
Db 1981 AATATCCGCGAGGCAAGGAGCGAGCCGAGAGGAGCTGGAGAAAGCTGCAGAACCGAGAG 2040  
Qy 2011 GATTTCTTGAAGGCAATCAGAAAGAGCTGTTGGAAGCTGAGGAAGCGCCGCTTCTCTG 2070  
Db 2041 GATTTCTTGAAGGCAATCAGAAAGAGCTGTTGGAAGCTGAGGAAGCGCCGCTTCTCTG 2100  
Qy 2071 GAGAAAGAGTAAAGAGCTAGAGCCATGGAGCCTAGAGAAACAGACTGAAGGATGAC 2130  
Db 2101 GAGAAAGAGTAAAGAGCTAGAGCCATGGAGCCTAGAGAAACAGACTGAAGGATGAC 2160  
Qy 2131 ATCCAGACAAATCCCAAGATCCAGCAGATGGCTGATTAATTTCTGGAGTCCAGAG 2190

Db 2161 ATCCAGACAAATCCCAAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2220  
Qy 2191 AAAACATCGGAGGCCCAAGTCTCAGCCAGAGCACTAGAGTGCACCTGGAACAGAAAGAGAG 2250  
Db 2221 AAACATCGGAGGCCCAAGTCTCAGCCAGAGCACTAGAGTGCACCTGGAACAGAAAGAGAG 2280  
Qy 2251 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATAAAAGAAAGAGCTGGCTGAC 2310  
Db 2281 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATAAAAGAAAGAGCTGGCTGAC 2340  
Qy 2311 AAGGAGACATCGGAGAAATGATCAGAGACA CAGAGGAGGAGGCCCATGAGAAAGGGCAAA 2370  
Db 2341 AAGGAGACATCGGAGAAATGATCAGAGACA CAGAGGAGGAGGCCCATGAGAAAGGGCAAA 2400  
Qy 2371 ATTCTCAGCAACAGAGGCGATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGAA 2430  
Db 2401 ATTCTCAGCAACAGAGGCGATGATCAATGCTATGATTCAGATCCAGATCAGATCCCTGAA 2460  
Qy 2431 CAGAGGATTTGGAACTGTCTGAAAGCCAAATAAATTTGCAGCAATAGCAGTCTTTTACC 2490  
Db 2461 CAGAGGATTTGGAACTGTCTGAAAGCCAAATAAATTTGCAGCAATAGCAGTCTTTTACC 2520  
Qy 2491 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550  
Db 2521 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580  
Qy 2551 CTGGAGACACAGGCTGGGAAAGTTGGAGGCCAGAAACCGAAACTTGGAGGAGCAGCTGGAG 2610  
Db 2581 CTGGAGACACAGGCTGGGAAAGTTGGAGGCCAGAAACCGAAACTTGGAGGAGCAGCTGGAG 2640  
Qy 2611 AAGATCAGCCACCAAGACCAAGTGA CAAAGATCGGCTGTGGAACCTGGAGACAAGATTG 2670  
Db 2641 AAGATCAGCCACCAAGACCAAGTGA CAAAGATCGGCTGTGGAACCTGGAGACAAGATTG 2700  
Qy 2671 CCGGAGGTCAAGTCTAGAGCAGGAGCAGAAACTGGAGCTCAAGGCCAGCTCACAAGAG 2730  
Db 2701 CCGGAGGTCAAGTCTAGAGCAGGAGCAGAAACTGGAGCTCAAGGCCAGCTCACAAGAG 2760  
Qy 2731 CTAAGTCTCTCTGAGGAGCGAGTCAAGTTCAGTTCAGCCCTGAGGCTGCAAGGCG 2790  
Db 2761 CTAAGTCTCTCTGAGGAGCGAGTCAAGTTCAGTTCAGCCCTGAGGCTGCAAGGCG 2820  
Qy 2791 GCCCTGAGAGCCAGCTTCCGCGGAGCAGACAGAGCTGGAAGAGACCACAGCAGAGCT 2850  
Db 2821 GCCCTGAGAGCCAGCTTCCGCGGAGCAGAGAGCTGGAAGAGACCACAGCAGAGCT 2880  
Qy 2851 GAAGAGGAGATCCAGGCACTCAGGCA CATAAGAGTGAATCCAGGCAATTTGATGCT 2910  
Db 2881 GAAGAGGAGATCCAGGCACTCAGGCA CATAAGAGTGAATCCAGGCAATTTGATGCT 2940  
Qy 2911 CTTGTTAAGCTGTACTGTAAATCAGAGCTTGGAGGAGCAGCTTAACAGCTGACCGAG 2970  
Db 2941 CTTGTTAAGCTGTACTGTAAATCAGAGCTTGGAGGAGCAGCTTAACAGCTGACCGAG 3000  
Qy 2971 GACAACCTGAATCTCAAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGTTCT 3030  
Db 3001 GACAACCTGAATCTCAAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGTTCT 3060  
Qy 3031 GGCCGCAACAGAGATTTGTAACCTGCGAAGTGAAGTGAACCATCTCCGCGCGGAGATC 3090  
Db 3061 GGCCGCAACAGAGATTTGTAACCTGCGAAGTGAAGTGAACCATCTCCGCGCGGAGATC 3120  
Qy 3091 ACGGAAACAGAGATGAGCTTACAGCCAGAAACCAAGATGAGGCTCTGAAGACCAAG 3150  
Db 3121 ACGGAAACAGAGATGAGCTTACAGCCAGAAACCAAGATGAGGCTCTGAAGACCAAG 3180  
Qy 3151 TGCAACATGCTGGAGGAAAGCTGATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3210  
Db 3181 TGCAACATGCTGGAGGAAAGCTGATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240  
Qy 3211 AAAGAGCGGAGTGGAGGCCCTGAGGAGCGCTCTGCGGTGATGAGAAATCCCAAGTTTGA 3270  
Db 3241 AAAGAGCGGAGTGGAGGCCCTGAGGAGCGCTCTGCGGTGATGAGAAATCCCAAGTTTGA 3300

3271 QY TGTGCGGTTCCGAGAGTGTGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGAGCC 3330  
3301 Db TGTGCGGTTCCGAGAGTGTGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGAGCC 3360  
3331 QY GATCAGCGGATCACCGAGTCTCCGAGGTGTGGAGCTGGGAGCTGGGAGGACACAAGGCT 3390  
3361 Db GATCAGCGGATCACCGAGTCTCCGAGGTGTGGAGCTGGGAGTGGGAGGACACAAGGCT 3420  
3391 QY GAGATTCTCGCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCTCTCT 3450  
3421 Db GAGATTCTCGCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCTCTCT 3480  
3451 QY GACAAAGCTCAATGACTGGAGAGAGCATGCTATGCTTGAATGAATGAATGCTCCGAGAGCTTA 3510  
3481 Db GACAAAGCTCAATGACTGGAGAGAGCATGCTATGCTTGAATGAATGAATGCTCCGAGAGCTTA 3540  
3511 QY CAGCGAAGCTGGAGACTGAACAGAGCTCAACAGAGGCTTCTGAAAGAGCAAGCCAAA 3570  
3541 Db CAGCGAAGCTGGAGACTGAACAGAGCTCAACAGAGGCTTCTGAAAGAGCAAGCCAAA 3600  
3571 QY TTACAGCAGCAGATGAGCTGAGAGAAATACATTTTCGCTGACTCAAGGACTGCAA 3630  
3601 Db TTACAGCAGCAGATGAGCTGAGAGAAATACATTTTCGCTGACTCAAGGACTGCAA 3660  
3631 QY GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTGAATGAGTATCAGCTG 3690  
3661 Db GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGTGAATGAGTATCAGCTG 3720  
3691 QY GAAAACATTACAGTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGCACTATTTCCTCAA 3750  
3721 Db GAAAACATTACAGTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGCACTATTTCCTCAA 3780  
3751 QY CAAACCAAATCTGATTTCTGAGCCAAATGAGCAACCTGCTTAAAGAAAAG 3810  
3781 Db CAAACCAAATCTGATTTCTGAGCCAAATGAGCAACCTGCTTAAAGAAAAG -- 3838  
3811 QY GGTTTATTTAGTCGAGGAAAGAGGACCTGCTTTACCCACACAGGTTCTCTGAGTAC 3870  
3839 Db -----AGGTTCTCTGAGTAC 3855  
3871 QY AATGAGCTGAAGTGCCTCGAGAGGAGAAAGTCTCGTGTGAGAGCTAGAGAAAGCC 3930  
3856 Db AATGAGCTGAAGTGCCTCGAGAGGAGAAAGTCTCGTGTGAGAGCTAGAGAAAGCC 3915  
3931 QY CTTCAGAAACCCGATCGAGTCCGGTCCGCCCGGAGAGAGCTGCCACCGCAAGCA 3990  
3916 Db CTTCAGAAACCCGATCGAGTCCGGTCCGCCCGGAGAGAGCTGCCACCGCAAGCA 3975  
3991 QY ACGGACCCACCAACCCATCCACGCCAGCCACCGGAGCAGCAGATCGCCATGTCGCGC 4050  
3976 Db ACGGACCCACCAACCCATCCACGCCAGCCACCGGAGCAGCAGATCGCCATGTCGCGC 4035  
4051 QY ATCGTGGGTGCGCAGAGCAGCCAGTGCATGAGCTGTGGCCCGCCGATCCAGC 4110  
4036 Db ATCGTGGGTGCGCAGAGCAGCCAGTGCATGAGCTGTGGCCCGCCGATCCAGC 4095  
4111 QY CGCAGAAAGAGTCTTCACTCCAGAGGAAATTTAGTGGCGCTTTAAGGAAACGATGAC 4170  
4096 Db CGCAGAAAGAGTCTTCACTCCAGAGGAAATTTAGTGGCGCTTTAAGGAAACGATGAC 4155  
4171 QY CACAAATTTCTCAGGATTCACGTAAGTGAACATCGAGCCACAAAGTGTGCTGTG 4230  
4156 Db CACAAATTTCTCAGGATTCACGTAAGTGAACATCGAGCCACAAAGTGTGCTGTG 4215  
4231 QY TGTCTGGATACCGTGACATTTGAGCCGAGGATCCAAATGCTCGAATGTCAGGTGATG 4290  
4216 Db TGTCTGGATACCGTGACATTTGAGCCGAGGATCCAAATGCTCGAATGTCAGGTGATG 4275  
4291 QY TGTCACCCCAAGTGTCCAGTCTTGGCAGCCACCTGGGGCTTGCTGCTGAATATGCC 4350  
4276 Db TGTCACCCCAAGTGTCCAGTCTTGGCAGCCACCTGGGGCTTGCTGCTGAATATGCC 4335

4351 QY ACAACATTACCGAGGCTTCTGCGGTGACAAATGAATGAATCTCCCGAGTCTCCAGACCAAG 4410  
4336 Db ACAACATTACCGAGGCTTCTGCGGTGACAAATGAATGAATCTCCCGAGTCTCCAGACCAAG 4395  
4411 QY GAGCCAGCAGCAGCTTGCACCTGGAAGGTGATGAAGTGCAGGAATAACAAACGA 4470  
4396 Db GAGCCAGCAGCAGCTTGCACCTGGAAGGTGATGAAGTGCAGGAATAACAAACGA 4455  
4471 QY GGACAGCAAGGCTGGGACAGGAAGTACATTTCTGGAGGATCAAAAGTCTCATTTAT 4530  
4456 Db GGACAGCAAGGCTGGGACAGGAAGTACATTTCTGGAGGATCAAAAGTCTCATTTAT 4515  
4531 QY GACAAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4590  
4516 Db GACAAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4575  
4591 QY GACGGGATGTATCTATTTCATGCTGCGTTGGTCTTCCGAATCCCAATACAGCCAAA 4650  
4576 Db GACGGGATGTATCTATTTCATGCTGCGTTGGTCTTCCGAATCCCAATACAGCCAAA 4635  
4651 QY GCA----- 4653  
4636 Db GCAGATCTCCATACATACATGAAAGTGAATCTACCCGACACCACTGCTGGCCCGGG 4695  
4654 Db ----- 4653  
4696 QY AGAACCTTACTTGTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGTCAACCGCTTA 4755  
4654 Db -----GAAAAGCAGAAAGCTGATTAACCTG 4680  
4756 QY GAATCAGTGTGCGCAGGTGGAGAGTCTTAGGGAAAAAGCAGAAAGTGTGCTAAACCTG 4815  
4681 Db CTTGGAAGCTCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAATGCAACCTG 4740  
4816 QY CTTGGAAGCTCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAATGCAACCTG 4875  
4741 Db CTTTTCAGTGACCAAGTGTGTTGTTGGGCAACGAGGAAGGCTCTACGCGCTGAATGTC 4800  
4876 QY CTTTTCAGTGACCAAGTGTGTTGTTGGGCAACGAGGAAGGCTCTACGCGCTGAATGTC 4935  
4801 Db TTGAAAACCTCCCTAAACCATGTCCAGGAATTTGGAGCAGCTCTTCCAAATTTATATTATC 4860  
4936 QY TTGAAAACCTCCCTAAACCATGTCCAGGAATTTGGAGCAGCTCTTCCAAATTTATATTATC 4995  
4861 Db AAGACCTTGAGAAAGTACTCATGATAGCAGGAGAGAGGCGGCACTGTCTTTGTGGAC 4920  
4996 QY AAGACCTTGAGAAAGTACTCATGATAGCAGGAGAGAGGCGGCACTGTCTTTGTGGAC 5055  
4921 Db GTGAGAAAGTGAACAGTCTTGGCCAGTCCCACTGCTGCGCCAGCCGAGCATCTCA 4980  
5056 QY GTGAGAAAGTGAACAGTCTTGGCCAGTCCCACTGCTGCGCCAGCCGAGCATCTCA 5115  
4981 Db CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTTGTTGGGCGAGCAAGATTGAGAAC 5040  
5116 QY CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTTGTTGGGCGAGCAAGATTGAGAAC 5175  
5041 Db GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTCTGCTATTTCTCCGCTCAACAGAAAC 5100  
5176 QY GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTCTGCTATTTCTCCGCTCAACAGAAAC 5235  
5101 Db CTCAGCAAAATCTGCATCTCGGAAAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5160  
5236 QY CTCAGCAAAATCTGCATCTCGGAAAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5295  
5161 Db TTCACCAATTAAGTATCTCTATTGGAACCAATTAATTTCTACGAAATCGCATGAAGCAG 5220  
5296 QY TTCACCAATTAAGTATCTCTATTGGAACCAATTAATTTCTACGAAATCGCATGAAGCAG 5355  
5221 Db TACAGCTCGAGGAATTCCTGGATAAGAAAGACCATTCCTTGGACCTGCTGTGTTGCC 5280  
5356 QY TACAGCTCGAGGAATTCCTGGATAAGAAAGACCATTCCTTGGACCTGCTGTGTTGCC 5415  
5281 Db GCCTCTTCCACAGCTTCCCTGTCTCAATCTGAGGTGAACAGCGGAGCGGAGAG 5340

Db 5416 GCCTCTTCCACACGCTTCCCTGTCTCAATCGTGCAGGTGAACACGCGCAGGCGACGAGAG 5475  
Qy 5341 GAGTACTTGTCTGTTTCCAGCAATTTGGAGTGTTCGTGGATCTTACGGAACACGTAGC 5400  
Db 5476 GAGTACTTGTCTGTTTCCAGCAATTTGGAGTGTTCGTGGATCTTACGGAACACGTAGC 5535  
Qy 5401 CGCACAGCATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGCTACAGAGAACCTTAT 5460  
Db 5536 CGCACAGCATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGCTACAGAGAACCTTAT 5595  
Qy 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTAAGATCCAGGACGCTCCTCAGCA 5520  
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTAAGATCCAGGACGCTCCTCAGCA 5655  
Qy 5521 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAACCCCGGCTACCTGGGCGCCCTGCCATT 5580  
Db 5656 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAACCCCGGCTACCTGGGCGCCCTGCCATT 5715  
Qy 5581 TCCTCAGAGCGATTTACTTTGGCGTCTCATACAGGATAAATTAAGGTTCATTCTGCTGC 5640  
Db 5716 TCCTCAGAGCGATTTACTTTGGCGTCTCATACAGGATAAATTAAGGTTCATTCTGCTGC 5775  
Qy 5641 AAGGGAAACCTCGTGAAGAGTCCGGCACTGAACACACCGGGGCGCCGCTCCACCTCCCGC 5700  
Db 5776 AAGGGAAACCTCGTGAAGAGTCCGGCACTGAACACACCGGGGCGCCGCTCCACCTCCCGC 5835  
Qy 5701 AGACGCCCCAAAGCGAGGCCCCACCCACGTAACAAGACATACCAAGCGGTGGCC 5760  
Db 5836 AGACGCCCCAAAGCGAGGCCCCACCCACGTAACAAGACATACCAAGCGGTGGCC 5895  
Qy 5761 TCAGCCGAGCGCGCCGCGGAGGCCCCAGCCACCGCGAGAGCAAGACACCCACCGC 5820  
Db 5896 TCAGCCGAGCGCGCGCGGAGGCCCCAGCCACCGCGAGAGCAAGACACCCACCGC 5955  
Qy 5821 TACCGGAGGGCGGACCGAGTGCAGAGGACAAAGTCTCTGCGCGCCCGCTCCGAGCGA 5880  
Db 5956 TACCGGAGGGCGGACCGAGTGCAGAGGACAAAGTCTCTGCGCGCCCGCTCCGAGCGA 6015  
Qy 5881 GAGAAGTCCCGCGCGAGTCTCAGACGCGAGAGCGGTCCCGCGGAGGCTGTTT 5940  
Db 6016 GAGAAGTCCCGCGCGAGTCTCAGACGCGAGAGCGGTCCCGCGGAGGCTGTTT 6075  
Qy 5941 GAAGACAGCAGCAGGGCGCGCTGCGTGGGAGCGGTGAGGACCCCGCTGCCAGGTG 6000  
Db 6076 GAAGACAGCAGCAGGGCGCGCTGCGTGGGAGCGGTGAGGACCCCGCTGCCAGGTG 6135  
Qy 6001 AACAGGGAGAGGGCAGAGTGC 6023  
Db 6136 AACAGGTCTGGGACCAAGTCTTC 6158

## RESULT 7

AX574425  
LOCUS AX574425  
DEFINITION Sequence 1 from Patent WO02059325.  
ACCESSION AX574425  
VERSION AX574425.1 GI:27551751

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Yu, X., Miranda, M. and Fridde, C.J.  
TITLE Human kinases and polynucleotides encoding the same  
JOURNAL Patent: WO 02059325-A 1 01-AUG-2002;  
Lexicon Genetics Incorporated (US)

FEATURES  
source  
1. .6165  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 86.1%; Score 5661.4; DB 6; Length 6165;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;  
Qy 19 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGTGAACCCATT 78  
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGTGAACCCATT 60  
Qy 79 GCCAGCCGGGCTCCAGGCTCAATCTCTTCCAGGGGAACACCCCTTTATGACTCAA 138  
Db 61 GCCAGCCGGGCTCCAGGCTCAATCTCTTCCAGGGGAACACCCCTTTATGACTCAA 120  
Qy 139 CAGCAGATGTCTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198  
Db 121 CAGCAGATGTCTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
Qy 199 GAATGCAGTCAAGCTGCTCTGATGAAGATTAAAGCAAGTGAAGCACTTTGTCGGAAGTAT 258  
Db 181 GAATGCAGTCAAGCTGCTCTGATGAAGATTAAAGCAAGTGAAGCACTTTGTCGGAAGTAT 240  
Qy 259 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGGACTTCGAAAGTCAGA 318  
Db 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGGACTTCGAAAGTCAGA 300  
Qy 319 AGTCTGTAGTGTGTGTCATTTGCTGAAGTCAGGTGTGAAGAGAAAGCAACCGGG 378  
Db 301 AGTCTGTAGTGTGTGTCATTTGCTGAAGTCAGGTGTGAAGAGAAAGCAACCGGG 360  
Qy 379 GACATCTATGCTATGAAGTGAAGCAAGGCTTTATTGGCCAGAGAGAGGTTTCA 438  
Db 361 GACATCTATGCTATGAAGTGAAGCAAGGCTTTATTGGCCAGAGAGAGGTTTCA 420  
Qy 439 TTTTGTGAGGAGCGGAAATATTATCTCGAAGCAACAAGCCCGTGGATCCCCCAATTA 498  
Db 421 TTTTGTGAGGAGCGGAAATATTATCTCGAAGCAACAAGCCCGTGGATCCCCCAATTA 480  
Qy 499 CAGTATGCTTTCAGGACCAAAATACCTTTATCTGATGAGCAATATCAGCTCGAGGG 558  
Db 481 CAGTATGCTTTCAGGACCAAAATACCTTTATCTGATGAGCAATATCAGCTCGAGGG 540  
Qy 559 GACTGTGTCACATTTTGAATAGATATGAGGACAGTTAGATGAAACCTGATACAGTTT 618  
Db 541 GACTGTGTCACATTTTGAATAGATATGAGGACAGTTAGATGAAACCTGATACAGTTT 600  
Qy 619 TACCTAGCTAGCTGATTTTGGCTGTTTCAAGGCTTCTGATGGATGATGATGATGATG 678  
Db 601 TACCTAGCTAGCTGATTTTGGCTGTTTCAAGGCTTCTGATGGATGATGATGATGATG 660  
Qy 679 GACATCAAGCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTGGATTT 738  
Db 661 GACATCAAGCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTGGATTT 720  
Qy 739 GGATCTCCCGGAAATGAATTCAAACAGATGGTGAATGCCAACTCCCGATGGGACC 798  
Db 721 GGATCTCCCGGAAATGAATTCAAACAGATGGTGAATGCCAACTCCCGATGGGACC 780  
Qy 799 CCAGATTACATGCTGCTGAGGCTGCTGATGAAAGGAGTGAAGGAGGAGGAGGAGGAG 858  
Db 781 CCAGATTACATGCTGCTGAGGCTGCTGATGAAAGGAGTGAAGGAGGAGGAGGAGGAG 840  
Qy 859 GGCTGTGACTGTGACTGCTGAGTGGGCTGATTCCTATGAGATGATTTATGGGAGA 918  
Db 841 GGCTGTGACTGTGACTGCTGAGTGGGCTGATTCCTATGAGATGATTTATGGGAGA 900  
Qy 919 TCCCGCTTCAGAGGGGACCTCTGCGCAAGACCTTCAATAACATTAATGATTTCCAGCG 978  
Db 901 TCCCGCTTCAGAGGGGACCTCTGCGCAAGACCTTCAATAACATTAATGATTTCCAGCG 960  
Qy 979 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1038  
Db 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020

QY 1039 TTGTTGTGGCCAGAAAAGAGAGACTGAAGTTTGAAGTCTTTGTGTCATCTCTTTCTTC 1098  
Db 1021 TTGTTGTGGCCAGAAAAGAGAGACTGAAGTTTGAAGTCTTTGTGTCATCTCTTTCTTC 1080  
QY 1099 TCTAAAATTGACTGGAAACAACTTGGTAACCTTCTCTCCCTTCTGTTCCACCTCAAG 1158  
Db 1081 TCTAAAATTGACTGGAAACAACTTGGTAACCTTCTCTCCCTTCTGTTCCACCTCAAG 1140  
QY 1159 TCTGACGATGACACCTTCCAAATTTGATGAACAGAGAGAATTCGTGGGTTTCATCTCT 1218  
Db 1141 TCTGACGATGACACCTTCCAAATTTGATGAACAGAGAGAATTCGTGGGTTTCATCTCT 1200  
QY 1219 CCGTGCCACTGAGCCCTCAGGCTTCTCGGTTGAGAACTCCCGTTTGTGGGTTTTCG 1278  
Db 1201 CCGTGCCACTGAGCCCTCAGGCTTCTCGGTTGAGAACTCCCGTTTGTGGGTTTTCG 1260  
QY 1279 TACAGCAAGCACTGGGATCTTGTGTAGATCTGAGTCTGTTGTGCGGTCTGACTCC 1338  
Db 1261 TACAGCAAGCACTGGGATCTTGTGTAGATCTGAGTCTGTTGTGCGGTCTGACTCC 1320  
QY 1339 CTTGCCAAGACTAGCTCCATGGAAGAAAATTTCTCATCAAAAGCAAGAGACTACAAGAC 1398  
Db 1321 CTTGCCAAGACTAGCTCCATGGAAGAAAATTTCTCATCAAAAGCAAGAGACTACAAGAC 1380  
QY 1399 TCTCAGCAAGTGTCAAGATGGAGCAGGAATGACCCGTTTACATCGGAGGTGCA 1458  
Db 1381 TCTCAGCAAGTGTCAAGATGGAGCAGGAATGACCCGTTTACATCGGAGGTGCA 1440  
QY 1459 GAGTGGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCCTCTCAGACTCAGAGA 1518  
Db 1441 GAGTGGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTGAAGCCCTCTCAGACTCAGAGA 1500  
QY 1519 TCCCTCTCGAGCAGACCTTGCTACTACATCAAGATGAGTAGCTTAAAGGGAAGT 1578  
Db 1501 TCCCTCTCGAGCAGACCTTGCTACTACATCAAGATGAGTAGCTTAAAGGGAAGT 1560  
QY 1579 TTGAGCAAGCAGGATGAGGTGTCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT 1638  
Db 1561 TTGAGCAAGCAGGATGAGGTGTCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT 1620  
QY 1639 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGAGTACAGGCTCAA 1698  
Db 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGAGTACAGGCTCAA 1680  
QY 1699 GTGGAGAAATGAGGTGATGATGAATCAGTTGGAGAGGATCTTCTCAGCAAGAAGA 1758  
Db 1681 GTGGAGAAATGAGGTGATGATGAATCAGTTGGAGAGGATCTTCTCAGCAAGAAGA 1740  
QY 1759 CGGAGTGATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGCTGAAGAAATTCAG 1818  
Db 1741 CGGAGTGATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGCTGAAGAAATTCAG 1800  
QY 1819 CGGAAGCGACAGAAATGACGATAAATCTTGAAGCTAAGATCAAGGAGGCTGAA 1878  
Db 1801 CGGAAGCGACAGAAATGACGATAAATCTTGAAGCTAAGATCAAGGAGGCTGAA 1860  
QY 1879 GTGGAGAAATGCGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938  
Db 1861 GTGGAGAAATGCGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
QY 1939 CTCAGAGAAATCGGAGAGGCT----- 1962  
Db 1921 CTCAGAGAAATCGGAGAGGCTGTAAGAACCCAGCAGCAGGAGCCACCGAGCTGTCGAG 1980  
QY 1963 -----GCARAGGAGCGAGCGAGGAGGCTGAGAGGCTGAGAAACCGAGAG 2010  
Db 1981 AATATCCCGCAGGCAAGAGGAGCGAGGAGGCTGGAGAGCTGCAGAAACCGAGAG 2040  
QY 2011 GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGGAACCGCCCAATCTCTG 2070  
Db 2041 GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGGAACCGCCCAATCTCTG 2100

QY 2071 GAGAAACAGGTAAGAGACTAGAGACCATGAGGCTAGAGAAACAGACTGAAAGGTGAC 2130  
Db 2101 GAGAAACAGGTAAGAGACTAGAGACCATGAGGCTAGAGAAACAGACTGAAAGGTGAC 2160  
QY 2131 ATCCAGACAAAATCCCAAAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2190  
Db 2161 ATCCAGACAAAATCCCAAAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2220  
QY 2191 AAACATCGGAGGCCCCAAGTCTCAGCCAGCACTTAGAAGTGCACCTGAAACAGAAAGAG 2250  
Db 2221 AAACATCGGAGGCCCCAAGTCTCAGCCAGCACTTAGAAGTGCACCTGAAACAGAAAGAG 2280  
QY 2251 CAGCACTATCAGGAAAAGATTAAAGTCTTGGAACAATCAGATAAAAGAAAGACCTGGCTGAC 2310  
Db 2281 CAGCACTATCAGGAAAAGATTAAAGTCTTGGAACAATCAGATAAAAGAAAGACCTGGCTGAC 2340  
QY 2311 AAGGAGACACTGGGAAACATGATCAGAGACAAGAGAGAGGCCCATGAGAGGGGCAAA 2370  
Db 2341 AAGGAGACACTGGGAAACATGATCAGAGACAAGAGAGAGGCCCATGAGAGGGGCAAA 2400  
QY 2371 ATTTCTCAGCAACAGAGGCGATGATCAATGCTATGATTCCAGATCAGATCCCTCGAA 2430  
Db 2401 ATTTCTCAGCAACAGAGGCGATGATCAATGCTATGATTCCAGATCAGATCCCTCGAA 2460  
QY 2431 CAGAGGATTTGTGAACTGTCTGAAAGCAATAAACTTGCAGCAATAGCAGCTTTTACC 2490  
Db 2461 CAGAGGATTTGTGAACTGTCTGAAAGCAATAAACTTGCAGCAATAGCAGCTTTTACC 2520  
QY 2491 CAAAGGAAATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550  
Db 2521 CAAAGGAAATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580  
QY 2551 CTGAGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAACTGAGGAGCAGCTGGAG 2610  
Db 2581 CTGAGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAACTGAGGAGCAGCTGGAG 2640  
QY 2611 AAGATCAGCCACCAAGACCAAGTGAACAAGTCTGGTCTGTGAACCTGGAGACAAGATTG 2670  
Db 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGTCTGGTCTGTGAACCTGGAGACAAGATTG 2700  
QY 2671 CGGAGGCTCAGCTTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCCGCAAGTCCACAGAG 2730  
Db 2701 CGGAGGCTCAGCTTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCCGCAAGTCCACAGAG 2760  
QY 2731 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCCCTGCAAGGCTGCAAGGCG 2790  
Db 2761 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCCCTGCAAGGCTGCAAGGCG 2820  
QY 2791 GCGCTGGAGCAGCTTCGCCAGGCGAAGCAGAGCTGGAGAGACCAAGCAGAGAGCT 2850  
Db 2821 GCGCTGGAGCAGCTTCGCCAGGCGAAGCAGAGCTGGAGAGACCAAGCAGAGAGCT 2880  
QY 2851 GAAGAGAGATCCAGGCACTCAGGCAATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2910  
Db 2881 GAAGAGAGATCCAGGCACTCAGGCAATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2940  
QY 2911 CTTGTAACAGCTGATCTGTAATCAGACTGAGAGAGCTGAAACAGCTGACCGAG 2970  
Db 2941 CTTGTAACAGCTGATCTGTAATCAGACTGAGAGAGCTGAAACAGCTGACCGAG 3000  
QY 2971 GACACGCTGAACCTCAACAAACAACTTCTACTTTGTCAAACAACTCGATGAGGCTTCT 3030  
Db 3001 GACACGCTGAACCTCAACAAACAACTTCTACTTTGTCAAACAACTCGATGAGGCTTCT 3060  
QY 3031 GCGCCAAACGAGAGATTGTAACAATCGAAAGTGAAGTGAAGTCCCGCGGAGATC 3090  
Db 3061 GCGCCAAACGAGAGATTGTAACAATCGAAAGTGAAGTGAAGTCCCGCGGAGATC 3120  
QY 3091 ACGGAACAGAGATGACGCTTACAGCCAGAAAGCAAGTGGAGGCTCTCAAGACCAAG 3150  
Db 3121 ACGGAACAGAGATGACGCTTACAGCCAGAAAGCAAGTGGAGGCTCTCAAGACCAAG 3180  
QY 3151 TGCACCATGCTGGAGGAACAGGCTCATGATTGGAGGCCCTTAAACGATGAGCTGCTAGAA 3210

Db 3181 TGCACCATGCTGGAGGAACAGGTGATGATTTGGAGGCCCTTAACAGGATGAGCTGCTAGAA 3240  
QY |||||  
Db 3211 AAAGAGCGGCAAGTGGAGGCTGAGGAGCGTCTCGGTGATGAGAAATCCAGTTTGAG 3270  
QY |||||  
Db 3241 AAAGAGCGGCAAGTGGAGGCTGAGGAGCGTCTCGGTGATGAGAAATCCAGTTTGAG 3300  
QY |||||  
Db 3271 TGTGGGTTCCAGAGCTGAGAGAAATGCTGGACACCGAGAAACAGAGACGAGGCGAGAGCC 3330  
QY |||||  
Db 3301 TGTGGGTTCCAGAGCTGAGAGAAATGCTGGACACCGAGAAACAGAGACGAGGCGAGAGCC 3360  
QY |||||  
Db 3331 GATCAGCGGATACCGAGTCTCGCCAGGTGGTGGAGCTGCAGTGAAGGACACAGGCT 3390  
QY |||||  
Db 3361 GATCAGCGGATACCGAGTCTCGCCAGGTGGTGGAGCTGCAGTGAAGGACACAGGCT 3420  
QY |||||  
Db 3391 GAGATTCTCGTCTGACGAGGCTCTCAAGAGCAGAGGCTGAAGGCCGAGAGGCTCTCT 3450  
Db 3421 GAGATTCTCGTCTGACGAGGCTCTCAAGAGCAGAGGCTGAAGGCCGAGAGGCTCTCT 3480  
QY |||||  
Db 3451 GACAGGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAGGCTTA 3510  
QY |||||  
Db 3481 GACAGGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAGGCTTA 3540  
QY |||||  
Db 3511 CAGCAGAGGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3570  
Db 3541 CAGCAGAGGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3600  
QY |||||  
Db 3571 TTACAGAGCAGATGAGCTGCGAGAAATCAATTTTCGGTCTGACTCAAGGACTGCAA 3630  
Db 3601 TTACAGAGCAGATGAGCTGCGAGAAATCAATTTTCGGTCTGACTCAAGGACTGCAA 3660  
QY |||||  
Db 3631 GAAGCTCTAGATCGGCTGATCTACTGAACAGAGAGAGTCACTTGGAGTATCAGCTG 3690  
QY |||||  
Db 3661 GAAGCTCTAGATCGGCTGATCTACTGAACAGAGAGAGTCACTTGGAGTATCAGCTG 3720  
QY |||||  
Db 3691 GAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGCACTATTTCTCAA 3750  
Db 3721 GAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGCACTATTTCTCAA 3780  
QY |||||  
Db 3751 CAACCAAACTCATGATTTCTGCAAGCAAAATGACCACTCTAAAGAGAAAG 3810  
Db 3781 CAACCAAACTCATGATTTCTGCAAGCAAAATGACCACTCTAAAGAGAAAG -- 3838  
QY |||||  
Db 3811 GGTATTATAGTCGACGGAAGAGGACCCCTGCTTTACCCACACAGGTTCTCTGCAGTAC 3870  
Db 3839 -----AGTTCTCTGCAGTAC 3855  
QY |||||  
Db 3871 AATGAGCTGAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGAGCC 3930  
Db 3856 AATGAGCTGAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGAGCC 3915  
QY |||||  
Db 3931 CTTTCAGAGACCCGATCGAGCTCCGCTCGCCGGGAGAGCTGCCACCGCAAGCA 3990  
Db 3916 CTTTCAGAGACCCGATCGAGCTCCGCTCGCCGGGAGAGCTGCCACCGCAAGCA 3975  
QY |||||  
Db 3991 ACAGACACCCACCCATCCAGCCAGCCACCGAGGAGCAGATGCCATGTCGCGC 4050  
Db 3976 ACAGACACCCACCCATCCAGCCAGCCACCGAGGAGCAGATGCCATGTCGCGC 4035  
QY |||||  
Db 4051 ATCGTGGGTCGCGAGACACAGCCAGTGCATGAGCCTGCTGGCCCGCCATCCAGC 4110  
Db 4036 ATCGTGGGTCGCGAGACACAGCCAGTGCATGAGCCTGCTGGCCCGCCATCCAGC 4095  
QY |||||  
Db 4111 CGCAGAAAGAGTCTTCAATCCAGAGGAAATTAGTCGCGCTTTAAGGAACGATGCAC 4170  
Db 4096 CGCAGAAAGAGTCTTCAATCCAGAGGAAATTAGTCGCGCTTTAAGGAACGATGCAC 4155  
QY |||||  
Db 4171 CACAATATTCTCACCGATTCAAGCTAGGACTGAACATCGAGGCGACAAAGTGTCTGTG 4230  
Db 4156 CACAATATTCTCACCGATTCAAGCTAGGACTGAACATCGAGGCGACAAAGTGTCTGTG 4215  
QY |||||  
Db 4231 TGTCTGGATACCGTGCATTTGAGCGCCAGGATCCAAATGCTCGAATGTCAAGTATG 4290

Db 4216 TGTCTGGATACCGTGCATTTGGAGCGCAGGATCAAAATGCTCGAATGTCAAGTATG 4275  
QY |||||  
Db 4291 TGTCAACCCCAAGTGTCCACGCTGCTTGCAGCACCTGCGCTTGCCTGTGAAATATGCC 4350  
QY |||||  
Db 4276 TGTCAACCCCAAGTGTCCACGCTGCTTGCAGCACCTGCGCTTGCCTGTGAAATATGCC 4335  
QY |||||  
Db 4351 ACACACTTCCACGAGGCTTCTGCGGTGACAAAATGAATGAACTCCAGAGTCTCCAGACCAAG 4410  
Db 4336 ACACACTTCCACGAGGCTTCTGCGGTGACAAAATGAATGAACTCCAGAGTCTCCAGACCAAG 4395  
QY |||||  
Db 4411 GAGCCACGAGCAGCTTGCACCTGGAAGGTGGAATGAAGTGCAGAGGATGAACAGCA 4470  
Db 4396 GAGCCACGAGCAGCTTGCACCTGGAAGGTGGAATGAAGTGCAGAGGATGAACAGCA 4455  
QY |||||  
Db 4471 GGACAGCAAGGCTGGGACAGGAATGATTTCTTGGAGGGATCAAAAGTCTCTCATTTAT 4530  
Db 4456 GGACAGCAAGGCTGGGACAGGAATGATTTCTTGGAGGGATCAAAAGTCTCTCATTTAT 4515  
QY |||||  
Db 4531 GACAAATGAAGCCAGAGAGCTGGACAGAGGCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4590  
Db 4516 GACAAATGAAGCCAGAGAGCTGGACAGAGGCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4575  
QY |||||  
Db 4591 GACGGGATGATCTATTTCATGCTGCTTCCGAACTCGCAAAATACAGGCCAAA 4650  
Db 4576 GACGGGATGATCTATTTCATGCTGCTTCCGAACTCGCAAAATACAGGCCAAA 4635  
QY |||||  
Db 4651 GCA----- 4653  
Db 4636 GCAGATGTCCCATACATACATGGAATCTCACCGGCACACACCTGCTGCGCCGGG 4695  
QY |||||  
Db 4654 ----- 4653  
Db 4696 AGAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGGTACCGCCTTA 4755  
QY |||||  
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAGAAAGCAGAGCTGATGCTAACTG 4815  
QY |||||  
Db 4691 CTTGGAACTCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAAGTGCAGCGCTG 4740  
Db 4816 CTTGGAACTCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAAGTGCAGCGCTG 4875  
QY |||||  
Db 4741 CCTTTCACTGACAGGCTGCTGCTGGGCAACGAGGAGGCTCTACGCGCTGAATGTC 4800  
Db 4876 CCTTTCACTGACAGGCTGCTGCTGGGCAACGAGGAGGCTCTACGCGCTGAATGTC 4935  
QY |||||  
Db 4801 TTGAAAACTCCCTAACCCATGTCCTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATC 4860  
Db 4936 TTGAAAACTCCCTAACCCATGTCCTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATC 4995  
QY |||||  
Db 4861 AAGCACTTGAGAGAGTCTCATGATAGCAGAGAGAGCGGCACTGTCTTTGTTGAC 4920  
Db 4996 AAGCACTTGAGAGAGTCTCATGATAGCAGAGAGAGCGGCACTGTCTTTGTTGAC 5055  
QY |||||  
Db 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCGCCGACATCTCA 4980  
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCGCCGACATCTCA 5115  
QY |||||  
Db 4981 CCCAACATTTTGAAGCTGCAAGGCTGCCACTGTTTGGGCGAGCAGATTTGAAC 5040  
Db 5116 CCCAACATTTTGAAGCTGCAAGGCTGCCACTGTTTGGGCGAGCAGATTTGAAC 5175  
QY |||||  
Db 5041 GGGCTCTGCACTGTGTCAGCCATGCCAGCAAAAGTCTCATTTCTCGCTACCAACGAAAC 5100  
Db 5176 GGGCTCTGCACTGTGTCAGCCATGCCAGCAAAAGTCTCATTTCTCGCTACCAACGAAAC 5235  
QY |||||  
Db 5101 CTGAGCAATTAAGTCTGCAATCCGGAAGAGATAGACCTCGAGCCCTGAGCTGTATCCAC 5160  
Db 5236 CTGAGCAATTAAGTCTGCAATCCGGAAGAGATAGACCTCGAGCCCTGAGCTGTATCCAC 5295  
QY |||||  
Db 5161 TTCAACAAATTAAGTCTGCAATTTGAAGCAATTAATTTCTAGAAATCGACATGAAGCAG 5220  
Db 5296 TTCAACAAATTAAGTCTGCAATTTGAAGCAATTAATTTCTAGAAATCGACATGAAGCAG 5355

QY 5221 TACAGCTCGAGGAATTCCTGGATAAGATGACCAATTCCTTGGACACCTGCTGTGTTGCC 5280  
Db 5356 TACAGCTCGAGGAATTCCTGGATAAGATGACCAATTCCTTGGACACCTGCTGTGTTGCC 5415  
QY 5281 GCCTCTTCCACAGCTTCCCTCTCAATCGTGCAGGTGAACAGCGCGAGGCGAGGAG 5340  
Db 5416 GCCTCTTCCACAGCTTCCCTCTCAATCGTGCAGGTGAACAGCGCGAGGCGAGGAG 5475  
QY 5341 GAGTACTGCTGCTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAGCGTAGC 5400  
Db 5476 GAGTACTGCTGCTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAGCGTAGC 5535  
QY 5401 CGACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGGCTTTGCCATACAGAAACCCCTAT 5460  
Db 5536 CGACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGGCTTTGCCATACAGAAACCCCTAT 5595  
QY 5461 CTGTTTGTGACCACTTCAACTCACTCGAAGTAATGAGATCCAGCGAGCTCTCAGCA 5520  
Db 5596 CTGTTTGTGACCACTTCAACTCACTCGAAGTAATGAGATCCAGCGAGCTCTCAGCA 5655  
QY 5521 GGGACCCCTGCGCGAGCGTACCTGGACATCCGAAACCCGCGTACCTGGGCGCTGCCAT 5580  
Db 5656 GGGACCCCTGCGCGAGCGTACCTGGACATCCGAAACCCGCGTACCTGGGCGCTGCCAT 5715  
QY 5581 TCCTCAGGACCAATTAATCTTGGCGTCCCTCATACAGGATAAATTAAGGTCATTTGCTGC 5640  
Db 5716 TCCTCAGGACCAATTAATCTTGGCGTCCCTCATACAGGATAAATTAAGGTCATTTGCTGC 5775  
QY 5641 AAGGGAACCTCGTGAAGAGTCCGGCACTGAACACCCGCGGCGCTCCACCTCCCGC 5700  
Db 5776 AAGGGAACCTCGTGAAGAGTCCGGCACTGAACACCCGCGGCGCTCCACCTCCCGC 5835  
QY 5701 AGCAGCCCAACAGCGAGGCCACCCAGTACAGTACAGAGACATCAACAGCGGTGCGC 5760  
Db 5836 AGCAGCCCAACAGCGAGGCCACCCAGTACAGTACAGAGACATCAACAGCGGTGCGC 5895  
QY 5761 TCCAGCCAGCGCCCGAGGCCCCAGCCAGCCGCGAGAGCCAGCACACCCACCGC 5820  
Db 5896 TCCAGCCAGCGCCCGAGGCCCCAGCCAGCCGCGAGAGCCAGCACACCCACCGC 5955  
QY 5821 TACCAGCGAGGCGGACCGAGTGGCGAGGCAAGTCTCTGGCGCGCCCTCGAGGCGA 5880  
Db 5956 TACCAGCGAGGCGGACCGAGTGGCGAGGCAAGTCTCTGGCGCGCCCTCGAGGCGA 6015  
QY 5881 GAGAGTCCCCCGCGGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTT 5940  
Db 6016 GAGAGTCCCCCGCGGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTT 6075  
QY 5941 GAAGACAGCAGCGGCGCGGCTGCTGCGGAGCGCGTAGGACCCCGCTGTCCAGGTG 6000  
Db 6076 GAAGACAGCAGCGGCGCGGCTGCTGCGGAGCGCGTAGGACCCCGCTGTCCAGGTG 6135  
QY 6001 AACAGGGAAGGCGAGAGTGC 6023  
Db 6136 AACAGGTCGGGACCAAGTCTTC 6158

RESULT 8  
AX166510  
LOCUS AX166510 6159 bp DNA linear PAT 22-JUN-2001  
DEFINITION Sequence 1 from Patent WO0138503.  
ACCESSION AX166510  
VERSION AX166510.1 GI:14546855  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Ploymann, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R.,  
Flanagan, P. and Clary, D.S.  
TITLE Novel human protein kinases and protein kinase-like enzymes

JOURNAL Patent: WO 0138503-A 1 31-MAY-2001;  
Sugen, Inc. (US)  
FEATURES Location/Qualifiers  
source 1..6159  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 86.0%; Score 5650.4; DB 6; Length 6159;  
Best Local Similarity 95.9%; Pred. No. 0;  
Matches 5943; Conservative 0; Mismatches 11; Indels 246; Gaps 4;  
QY 19 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATCCTGGTCTGTGTAACCCCAT 78  
Db 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATCCTGGTCTGTGTAACCCCAT 60  
QY 79 GCACGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 138  
Db 61 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120  
QY 139 CAGCAGATGTCTCTCTTTCCGAGAAGGATATATAGATCCCTCTTTTGTCTCTTTGAA 198  
Db 121 CAGCAGATGTCTCTCTTTCCGAGAAGGATATATAGATCCCTCTTTTGTCTCTTTGAA 180  
QY 199 GAATGCACTCAGCTGCTCTGATGAGATTAAAGCACTGAGCAACTTTGTCGGA--G 255  
Db 181 GAATGCACTCAGCTGCTCTGATGAGATTAAAGCACTGAGCAACTTTTCCCGAAGTG 240  
QY 256 TATTCACACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGCGCAAGGACTTTCGAAGTC 315  
Db 241 TATTCACACACCATAGCTGAGTTACAGGAGCTTCGCGCAAGGACTTTCGAAGTC 300  
QY 316 AGAAGTCTTGTAGTTGTGTCTCTTGTGAGTGCAGTGTGTAGAGAGAAGCAACC 375  
Db 301 AGAAGTCTTGTAGTTGTGTCTCTTGTGAGTGCAGTGTGTAGAGAGAAGCAACC 360  
QY 376 GGGACATCTATGCTATGAAAGTATGAAGAAGGCTTTATTTGGCCACGAGCAGGTT 435  
Db 361 GGGACATCTATGCTATGAAAGTATGAAGAAGGCTTTATTTGGCCACGAGCAGGTT 420  
QY 436 TCAATTTTTCAGGAAGCGGAAACATATTATTCGAAGCAACAGCCCGTGGATCCCCCAA 495  
Db 421 TCAATTTTTCAGGAAGCGGAAACATATTATTCGAAGCAACAGCCCGTGGATCCCCCAA 480  
QY 496 TTACAGTATGCTTTCAGGACAAATATCCTTTATCTGATGGAGGAATATCAGCTGGA 555  
Db 481 TTACAGTATGCTTTCAGGACAAATATCCTTTATCTGATGGAGGAATATCAGCTGGA 540  
QY 556 GGGACCTGTGTCACTTTTGAATAGATATGAGGACCCAGTATGATAAACCCTGATACAG 615  
Db 541 GGGACCTGTGTCACTTTTGAATAGATATGAGGACCCAGTATGATAAACCCTGATACAG 600  
QY 616 TTTTACCTAGCTAGCTGATTTTGGCTGTTTCAGAGGTTTCTCTGATGGGATACGTGCAT 675  
Db 601 TTTTACCTAGCTAGCTGATTTTGGCTGTTTTCAGAGGTTTCTCTGATGGGATACGTGCAT 660  
QY 676 CGACATCAAGCTCGAAGCAATTCCTGTTGACCGCACAGGACACATCAAGCTGTTGGAT 735  
Db 661 CGACATCAAGCTCGAAGCAATTCCTGTTGACCGCACAGGACACATCAAGCTGTTGGAT 720  
QY 736 TTTGGATCTCCCGGAAATGAATTCAAACAAGATGTGAATGCCAAATCCCGATTGGG 795  
Db 721 TTTGGATCTCCCGGAAATGAATTCAAACAAGATGTGAATGCCAAATCCCGATTGGG 780  
QY 796 ACCCAGATACATGGCTCCTGAGTGTCTGATGAAACGGGATGAAAGGACCC 855  
Db 781 ACCCAGATACATGGCTCCTGAGTGTCTGATGAAACGGGATGAAAGGACCC 840  
QY 856 TACGGCTGAGCTGTGACTGCTGAGTGGCGGTGATTCCTTATGAGATGATTTATGGG 915  
Db 841 TACGGCTGAGCTGTGACTGCTGAGTGGCGGTGATTCCTTATGAGATGATTTATGGG 900

QY 916 AGATCCCCCTTCGAGAGGGAACCTCTGCCAGAAACCTTCAATTAACATTATGAATTTCCAG 975  
Db 901 AGATCCCCCTTCGAGAGGGAACCTCTGCCAGAAACCTTCAATTAACATTATGAATTTCCAG 960  
QY 976 CGGTTTTTGAAATTTCCAGATGACCCCAAGTAGAGAGTACGACCTTTCTGATCTGATTCAA 1035  
Db 961 CGGTTTTTGAAATTTCCAGATGACCCCAAGTAGAGAGTACGACCTTTCTGATCTGATTCAA 1020  
QY 1036 AGCTTGTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTTCTTCTGCCATCTTTTC 1095  
Db 1021 AGCTTGTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTTCTTCTGCCATCTTTTC 1080  
QY 1096 TTCTCTAATTAATGACTGGACACATTCGTAACCTCTCCCTCCCGCTCTGTTCCACCCCTC 1155  
Db 1081 TTCTCTAATTAATGACTGGACACATTCGTAACCTCTCCCTCCCGCTCTGTTCCACCCCTC 1140  
QY 1156 AAGTCTGACGATGACACCTCCAATTTTGTATGAACCCAGAGAAATTCGTGGGTTTCATCC 1215  
Db 1141 AAGTCTGACGATGACACCTCCAATTTTGTATGAACCCAGAGAAATTCGTGGGTTTCATCC 1200  
QY 1216 TCTCCGTGCCAGTGAACCCCTCAGGCTTCTCGGTGAAGAACTGCGGTTTGTGGGTTTT 1275  
Db 1201 TCTCCGTGCCAGTGAACCCCTCAGGCTTCTCGGTGAAGAACTGCGGTTTGTGGGTTTT 1260  
QY 1276 TCGTACAGCAAGGACCTGGGGATTTCTGGTAGATCTTGGTCTGTTGTTGGGTCTGGAC 1335  
Db 1261 TCGTACAGCAAGGACCTGGGGATTTCTGGTAGATCTTGGTCTGTTGTTGGGTCTGGAC 1320  
QY 1336 TCCCTGCCAAGACTAGCTCCATGGAAAAAACTTCTCATAAAAAGAAAGAGCTACAA 1395  
Db 1321 TCCCTGCCAAGACTAGCTCCATGGAAAAAACTTCTCATAAAAAGAAAGAGCTACAA 1380  
QY 1396 GACTCTCAGACAAAGTGTCAAGATGGAGCGAGAAATGACCCGGTTTACATCGGAGTG 1455  
Db 1381 GACTCTCAGACAAAGTGTCAAGATGGAGCGAGAAATGACCCGGTTTACATCGGAGTG 1440  
QY 1456 TCAGAGTGGAGCTGTGCTTAGTCAGAGGAGGTGGAGTGAAGCCCTCTGAGACTCAG 1515  
Db 1441 TCAGAGTGGAGCTGTGCTTAGTCAGAGGAGGTGGAGTGAAGCCCTCTGAGACTCAG 1500  
QY 1516 AGATCCCTCTGAGAGGAGCTTGTACTATCATCAGAAATGACAGTAAAGCGA 1575  
Db 1501 AGATCCCTCTGAGAGGAGCTTGTACTATCATCAGAAATGACAGTAAAGCGA 1560  
QY 1576 AGTTTGAGCAAGCAGGATGAGGTTGTCCAGAGGATGACAAAGCACTGACGCTTCTC 1635  
Db 1561 AGTTTGAGCAAGCAGGATGAGGTTGTCCAGAGGATGACAAAGCACTGACGCTTCTC 1620  
QY 1636 CATGATATCAGAGAGCAGCCGAGCTTCAAGAAATCAAGAGCAGGAGTACAGGT 1695  
Db 1621 CATGATATCAGAGAGCAGCCGAGCTTCAAGAAATCAAGAGCAGGAGTACAGGT 1680  
QY 1696 CAAGTGAAGAAATGAGTTGATGAATCACTTGGAAAGAGGATCTTCTCTCAGCAAGA 1755  
Db 1681 CAAGTGAAGAAATGAGTTGATGAATCACTTGGAAAGAGGATCTTCTCTCAGCAAGA 1740  
QY 1756 AGACGAGTGAATCTTACGAATCTGAGCTGAGAGTCTCGGTTGCTGCTGAAGATTC 1815  
Db 1741 AGACGAGTGAATCTTACGAATCTGAGCTGAGAGTCTCGGTTGCTGCTGAAGATTC 1800  
QY 1816 AAGCGAAACGACAGATGTACAGTAACTTGTGAAGCTTAAGGATCAAGGAGGCT 1875  
Db 1801 AAGCGAAACGACAGATGTACAGTAACTTGTGAAGCTTAAGGATCAAGGAGGCT 1860  
QY 1876 GAAGTGGGAAATATCGAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAG 1935  
Db 1861 GAAGTGGGAAATATCGAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAG 1920  
QY 1936 GAGCTCCAGAGAACTGGAGAGGCT----- 1962  
Db 1921 GAGCTCCAGAGAACTGGAGAGGCTGTAAAAGCAGCAGGAGGCTACCCAGCTGTG 1980  
QY 1963 -----GCAAGGAGCGAGCGGAGGAGCTGGAGAGCTGGAGAACTGAGAACCGA 2007

Db 1981 CAGAATATCCGCCAGGCAAGAGCGAGCCGAGAGGAGCTGGAGAAAGCTTCAGAACCGA 2040  
QY 2008 GAGGATTTCTTCTGAAGGATCAGAAAGAAAGCTGTGGAAGCTGAGGAACGCGCATTTCT 2067  
Db 2041 GAGGATTTCTTCTGAAGGATCAGAAAGAAAGCTGTGGAAGCTGAGGAACGCGCATTTCT 2100  
QY 2068 CTGGAGAAACAGGTAAGAGACTAGAGACCATGAGCGGTAGAGAAAAACAGACTGAAGAT 2127  
Db 2101 CTGGAGAAACAGGTAAGAGACTAGAGACCATGAGCGGTAGAGAAAAACAGACTGAAGAT 2160  
QY 2128 GACTCCAGACAAATCCCAACAGATCCAGCAGATGCTGATAAATTTCTGGAGCTCGAA 2187  
Db 2161 GACTCCAGACAAATCCCAACAGATCCAGCAGATGCTGATAAATTTCTGGAGCTCGAA 2220  
QY 2188 GAGAAACATCGGAGGCGCCCAAGTCTCAGCCAGCACCTAGAGTGCACCTGAAACAGAAA 2247  
Db 2221 GAGAAACATCGGAGGCGCCCAAGTCTCAGCCAGCACCTAGAGTGCACCTGAAACAGAAA 2280  
QY 2248 GAGCAGCACTATGAGGAAAGATTAAGTCTTGGCAATCAGATATAAGAGACCTGGCT 2307  
Db 2281 GAGCAGCACTATGAGGAAAGATTAAGTCTTGGCAATCAGATATAAGAGACCTGGCT 2340  
QY 2308 GACAAGGAGACACTGGAGAACATGATCAGAGACACAGAGAGGAGGCCCATGAGAGGCG 2367  
Db 2341 GACAAGGAGACACTGGAGAACATGATCAGAGACACAGAGAGGAGGCCCATGAGAGGCG 2400  
QY 2368 AAAATTCAGCGAAACAGAGGCGATGATCAATGCTATGATTTCCAAGATCAGATCCCTG 2427  
Db 2401 AAAATTCAGCGAAACAGAGGCGATGATCAATGCTATGATTTCCAAGATCAGATCCCTG 2460  
QY 2428 GAAACAGAGGATTTGGAACTGTCTGAAAGCAATAAATCTGCAGCAATAGCAGTCTTTT 2487  
Db 2461 GAAACAGAGGATTTGGAACTGTCTGAAAGCAATAAATCTGCAGCAATAGCAGTCTTTT 2520  
QY 2488 ACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTT 2547  
Db 2521 ACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTT 2580  
QY 2548 TACCTGAGACACAGGCTGGAGTTGGAGCCCGAGAACCGAAACTGGAGGAGCAGCTG 2607  
Db 2581 TACCTGAGACACAGGCTGGAGTTGGAGCCCGAGAACCGAAACTGGAGGAGCAGCTG 2640  
QY 2608 GAGAAGATCAAGCAACCAAGACCAAGTCAAGAAATCGGCTGTGGAACCTGGAGACAAGA 2667  
Db 2641 GAGAAGATCAAGCAACCAAGACCAAGTCAAGAAATCGGCTGTGGAACCTGGAGACAAGA 2700  
QY 2668 TTGCGGAGTCACTCTAGAGCAGGAGAGAGAACTGAGCTCAAGCGCCAGCTCACA 2727  
Db 2701 TTGCGGAGTCACTCTAGAGCAGGAGAGAGAACTGAGCTCAAGCGCCAGCTCACA 2760  
QY 2728 GAGCTACAGCTCTCCCTGACAGGCGCGAGTCAAGTTGACAGCCCTGCAAGGTTGACGG 2787  
Db 2761 GAGCTACAGCTCTCCCTGACAGGCGCGAGTCAAGTTGACAGCCCTGCAAGGTTGACGG 2820  
QY 2788 CGGCGCTTGAGAGCCAGCTTTCGCGAGCGAAGACAGAGCTGGAAGACCAAGAGAA 2847  
Db 2821 CGGCGCTTGAGAGCCAGCTTTCGCGAGCGAAGACAGAGCTGGAAGACCAAGAGAA 2880  
QY 2848 GCTGAAGAGAGATTCAGGCACTCACGGCACAATAGAGTGAATTCAGCGCCAAATTTGAT 2907  
Db 2881 GCTGAAGAGAGATTCAGGCACTCACGGCACAATAGAGTGAATTCAGCGCCAAATTTGAT 2940  
QY 2908 GCTTTCTGAAGCTGTAATCTGTAATCAGAGCTGAGAGGAGGAGTAAACAGAGTGAAC 2967  
Db 2941 GCTTTCTGAAGCTGTAATCTGTAATCAGAGCTGAGAGGAGGAGTAAACAGAGTGAAC 3000  
QY 2968 GAGGACAACTGAAGCTCAACAAACAACTTCTACTTTGCCAAAACAACTCGATGAGGCT 3027  
Db 3001 GAGGACAACTGAAGCTCAACAAACAACTTCTACTTTGCCAAAACAACTCGATGAGGCT 3060  
QY 3028 TCTGGCCCAACGACGAGATTGTACAACTCGAAGTGAAGTGAACCATTTCTCCCGGGAG 3087

Db 3061 TCTGGCGCAACGACGAGATTGTACAACTGCGAAGTGAAAGTGACCATCTCTCGCGGGAG 3120  
QY 3088 ATCAAGGAACGAGATGATGAGCTTACAGCCAGAGCAAAACCATGGAGCTCTGAAGACC 3147  
Db 3121 ATCAGGAACGAGATGATGAGCTTACAGCCAGAGCAAAACCATGGAGCTCTGAAGACC 3180  
QY 3148 AGTGACCAATGCTGGAGGAAACAGGTCTATGGAATTTGGAGGCCCTAAACGATGAGTGCTA 3207  
Db 3181 AGTGACCAATGCTGGAGGAAACAGGTCTATGGAATTTGGAGGCCCTAAACGATGAGTGCTA 3240  
QY 3208 GAAAAAGAGCGCAGTGGAGGCCCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGATTT 3267  
Db 3241 GAAAAAGAGCGCAGTGGAGGCCCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGATTT 3300  
QY 3268 GAGTGTGGGTTCGAGAGCTGAGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGA 3327  
Db 3301 GAGTGTGGGTTCGAGAGCTGAGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGA 3360  
QY 3328 GCCGATCAGCGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGACACAAG 3387  
Db 3361 GCCGATCAGCGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGACACAAG 3420  
QY 3388 GGTGAGATCTCGCTCTGCAGCAGGCTCTCAAGAGACAGAACTCAAGGCCGAGAGCCTC 3447  
Db 3421 GGTGAGATCTCGCTCTGCAGCAGGCTCTCAAGAGACAGAACTCAAGGCCGAGAGCCTC 3480  
QY 3448 TCTGACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAAGC 3507  
Db 3481 TCTGACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAAGC 3540  
QY 3508 TTACAGCAGAACTGGAGACTCAACGAGAGCTCAACAGAGAGCTCTGGAAGAGCAGGCC 3567  
Db 3541 TTACAGCAGAACTGGAGACTCAACGAGAGCTCAACAGAGAGCTCTGGAAGAGCAGGCC 3600  
QY 3568 AAATTACAGCAGATGGAACCTGAGAAATATCATATTTCCGCTCACTCAAGGACTG 3627  
Db 3601 AAATTACAGCAGATGGAACCTGAGAAATATCATATTTCCGCTCACTCAAGGACTG 3660  
QY 3628 GAAGAGCTTAGATCGGGCTGATCTACTGAGCAGAAAGAGTGGAGTATCAG 3687  
Db 3661 GAAGAGCTTAGATCGGGCTGATCTACTGAGCAGAAAGAGTGGAGTATCAG 3720  
QY 3688 CTGGAACCAATTCAGGTTCTCTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCT 3747  
Db 3721 CTGGAACCAATTCAGGTTCTCTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCT 3780  
QY 3748 CAACAAACCAACTCATTTGATTTCTGACGCCAAATGGACCAACCTGCTAAAGAA 3807  
Db 3781 CAACAAACCAACTCATTTGATTTCTGACGCCAAATGGACCAACCTGCTAAAGAA 3840  
QY 3808 AAGGGTTTATTTAGTCGACGGAAGAGGACCTGCTTTACCCACACAGGTTCCCTGCGAG 3867  
Db 3841 A-----AGTTCTCTGCG 3855  
QY 3868 TACAATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGTCTGCTGTCAGACTAGAGAA 3927  
Db 3856 TACAATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGTCTGCTGTCAGACTAGAGAA 3915  
QY 3928 GCCCTTCAGAGACCGCATCGAGCTCCGCTCCGCGGAGGAGCTGCCACCGCAA 3987  
Db 3916 GCCCTTCAGAGACCGCATCGAGCTCCGCTCCGCGGAGGAGCTGCCACCGCAA 3975  
QY 3988 GOACGGACCCACCCATCCACGCCAGCCACCGGAGGAGCAGATGCCATGTCC 4047  
Db 3976 GOACGGACCCACCCATCCACGCCAGCCACCGGAGGAGCAGATGCCATGTCC 4035  
QY 4048 GCATCTGTCGCTGCCAGAGACACAGCCAGTGCATGAGCTGTCGCCCGGCATCC 4107  
Db 4036 GCATCTGTCGCTGCCAGAGACACAGCCAGTGCATGAGCTGTCGCCCGGCATCC 4095  
QY 4108 AGCCGAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGTCTTAAGGAACGATG 4167  
Db 4096 AGCCGAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGTCTTAAGGAACGATG 4155

QY 4168 CACCACAAATTTCTCTCAACGATTCAACGTAGGACTGAACATGCGAGCCACAAAGTGTCT 4227  
Db 4156 CACCACAAATTTCTCTCAACGATTCAACGTAGGACTGAACATGCGAGCCACAAAGTGTCT 4215  
QY 4228 GTGTGTCTGATACCGTGCACTTTGGAGCCAGGCATCCAAATGTCTCGAATGTCAAGT 4287  
Db 4216 GTGTGTCTGATACCGTGCACTTTGGAGCCAGGCATCCAAATGTCTCGAATGTCAAGT 4275  
QY 4288 ATGTGTCAACCCCAAGTCTCCACGTCTTGCAGCCACCTGCGGCTTGCCTGTGCTGAAT 4347  
Db 4276 ATGTGTCAACCCCAAGTCTCCACGTCTTGCAGCCACCTGCGGCTTGCCTGTGCTGAAT 4335  
QY 4348 GCCACACATTTACCGAGGCTTTCTGCCGTGACAAATGAATCCACAGTCTCCAGACC 4407  
Db 4336 GCCACACATTTACCGAGGCTTTCTGCCGTGACAAATGAATCCACAGTCTCCAGACC 4395  
QY 4408 AAGAGCCACGACGAGCTTGCACCTTGGAGGCTGGATGAAGTGCAGGATTAACAA 4467  
Db 4396 AAGAGCCACGACGAGCTTGCACCTTGGAGGCTGGATGAAGTGCAGGATTAACAA 4455  
QY 4468 CGAGGACAGCAAGGCTGGGACAGGAATCATTTCTCTGAGGGATCAAAAGTCTCAT 4527  
Db 4456 CGAGGACAGCAAGGCTGGGACAGGAATCATTTCTCTGAGGGATCAAAAGTCTCAT 4515  
QY 4528 TATGACAAATGAAGCCAGAGAGCTGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTT 4587  
Db 4516 TATGACAAATGAAGCCAGAGAGCTGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTT 4575  
QY 4588 CCGACCGGATGTATCTTATTCATTTGTTGCTTCCAACTTCGCAATACAGCC 4647  
Db 4576 CCGACCGGATGTATCTTATTCATTTGTTGCTTCCAACTTCGCAATACAGCC 4635  
QY 4648 AAAGCAGA----- 4655  
Db 4636 AAAGCAGATGTCCTATACATACATACTCAAGATGGAATCTCACCCGACACCACTGTGCGCC 4695  
QY 4656 ----- 4655  
Db 4696 GGGAGAACCTTACTTTGCTAGCTCCAGCTTCCCTGAACAAACAGCGCTGGGTCAAGCC 4755  
QY 4656 -----AAAGCAGAAAGCTGATGCTAAA 4677  
Db 4756 TTAGATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAGAAACAGAGCTGATGCTAAA 4815  
QY 4678 CTGCTTGGAACTCCCTGCTGAAACTGGAAGTGTATGACGCTAGACATGATGACG 4737  
Db 4816 CTGCTTGGAACTCCCTGCTGAAACTGGAAGTGTATGACGCTAGACATGAACTGACG 4875  
QY 4738 CTGCTTCACTGACACAGGTGGTGTGGTGGGACCGAGGAGGCTCTACGCGCTGAAT 4797  
Db 4876 CTGCTTCACTGACACAGGTGGTGTGGTGGGACCGAGGAGGCTCTACGCGCTGAAT 4935  
QY 4798 GTCTTGAAGAACTCCCTAACCCATGTCCAGGAATTTGGAGAGCTTTCCAAATTTATTT 4857  
Db 4936 GTCTTGAAGAACTCCCTAACCCATGTCCAGGAATTTGGAGAGCTTTCCAAATTTATTT 4995  
QY 4858 ATCAAGGACTGAGAGCTTACTCATGATAGCAGGAGAGAGCGGACCTGCTCTGTG 4917  
Db 4996 ATCAAGGACTGAGAGCTTACTCATGATAGCAGGAGAGAGCGGACCTGCTCTGTG 5055  
QY 4918 GAGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATC 4977  
Db 5056 GAGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATC 5115  
QY 4978 TCACCCAAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGGGAGGAGGATGAG 5037  
Db 5116 TCACCCAAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGGGAGGAGGATGAG 5175  
QY 5038 AAGCGGCTTGCATCTGTGACGCAATGCCAGCAAGTCTCATTTCTCGCTACACGAA 5097  
Db 5176 AAGCGGCTTGCATCTGTGACGCAATGCCAGCAAGTCTCATTTCTCGCTACACGAA 5235

QY 5098 AACCTCAGCAATACATGATCCGGAAGAGATAGAGACCTCAGAGCCCTGACGTGTATC 5157  
Db |||||||  
QY 5236 AACCTCAGCAATACATGATCCGGAAGAGATAGAGACCTCAGAGCCCTGACGTGTATC 5295  
Db |||||||  
QY 5158 CACTTCACCAATACAGTATCTCTCATTTGGAACCAATAATTTCTAGAAATCGACATGAAG 5217  
Db |||||||  
QY 5296 CACTTCACCAATACAGTATCTCTCATTTGGAACCAATAATTTCTAGAAATCGACATGAAG 5355  
Db |||||||  
QY 5218 CAGTACAGCTCAGGAATTCCTGGATAGATGACCATTCCTTGGACCTCTGCTGTTT 5277  
Db |||||||  
QY 5356 CAGTACAGCTCAGGAATTCCTGGATAGATGACCATTCCTTGGACCTCTGCTGTTT 5415  
Db |||||||  
QY 5278 GCGGCTCTTCCAAACAGCTTCCCTGTCCTCAATCGTCAGGTGAACAGAGCCGAGGAGGGA 5337  
Db |||||||  
QY 5416 GCGGCTCTTCCAAACAGCTTCCCTGTCCTCAATCGTCAGGTGAACAGAGCCGAGGAGGGA 5475  
Db |||||||  
QY 5338 GAGAGTACTGTTGTTTCCAGAAATTTGGAGTGTCTGCGATTTCTTACGGAAGAGGT 5397  
Db |||||||  
QY 5476 GAGAGTACTGTTGTTTCCAGAAATTTGGAGTGTCTGCGATTTCTTACGGAAGAGGT 5535  
Db |||||||  
QY 5398 AGCGGACAGACGATCTCAAGTGGAGTCTGCTTACCTTTGGCCCTTGGCTTACAGAGAACCC 5457  
Db |||||||  
QY 5536 AGCGGACAGACGATCTCAAGTGGAGTCTGCTTACCTTTGGCCCTTGGCTTACAGAGAACCC 5595  
Db |||||||  
QY 5458 TATCTGTTTGTGACCACTTCAACTCACTCGAATTAATTGAGATCCAGGACGCTCTCTCA 5517  
Db |||||||  
QY 5596 TATCTGTTTGTGACCACTTCAACTCACTCGAATTAATTGAGATCCAGGACGCTCTCTCA 5655  
Db |||||||  
QY 5518 GCAGGACCCCTCCGAGGTACCTGGACATCCCGAACCCCGGCTACCTGGGCCCTGCC 5577  
Db |||||||  
QY 5656 GCAGGACCCCTCCGAGGTACCTGGACATCCCGAACCCCGGCTACCTGGGCCCTGCC 5715  
Db |||||||  
QY 5578 ATTTCCTCAGAGCGGATTTACTTTGGGCTCTCATACCGAGTAATAATTAGGGTCAATTGTC 5637  
Db |||||||  
QY 5716 ATTTCCTCAGAGCGGATTTACTTTGGGCTCTCATACCGAGTAATAATTAGGGTCAATTGTC 5775  
Db |||||||  
QY 5638 TGAAGGAAACCTCTGTAAGAGTCTGGGACATGAACACCCCGGGCCCGTCCACCTCC 5697  
Db |||||||  
QY 5776 TGAAGGAAACCTCTGTAAGAGTCTGGGACATGAACACCCCGGGCCCGTCCACCTCC 5835  
Db |||||||  
QY 5698 CGCAGAGCCCCAACAGAGAGCCCAACCCACAGTACCAAGAGCACATCAACAGGCGCGTG 5757  
Db |||||||  
QY 5836 CGCAGAGCCCCAACAGAGAGCCCAACCCACAGTACCAAGAGCACATCAACAGGCGCGTG 5895  
Db |||||||  
QY 5758 GCCTCCAGCCAGCGGCGCCGGAAGCCCGAGCCACCCCGGAGAGCCAGACACCCAC 5817  
Db |||||||  
QY 5896 GCCTCCAGCCAGCGGCGCCGGAAGCCCGAGCCACCCCGGAGAGCCAGACACCCAC 5955  
Db |||||||  
QY 5818 CGCTACCGGAGGGGGACCGAGCTGCGCAGGGAAGTCTCTGGCCCGCCCTGGAG 5877  
Db |||||||  
QY 5956 CGTACCGGAGGGGGACCGAGCTGCGCAGGGAAGTCTCTGGCCCGCCCTGGAG 6015  
Db |||||||  
QY 5878 CGAGAGAGTCCCGGCGCGATGCTCAGCAACCGGAGAGAGCGGTCCCGGAGGCGTG 5937  
Db |||||||  
QY 6016 CGAGAGAGTCCCGGCGCGATGCTCAGCAACCGGAGAGAGCGGTCCCGGAGGCGTG 6075  
Db |||||||  
QY 5938 TTTTGAAGACAGCAGGCGCGCTGCTGCGGAGCCGCTGAGGACCCCGCTGCTCCAG 5997  
Db |||||||  
QY 6076 TTTTGAAGACAGCAGGCGCGCTGCTGCGGAGCCGCTGAGGACCCCGCTGCTCCAG 6135  
Db |||||||  
QY 5998 GTGAACAAGGGAAGAGGGCA 6017  
Db |||||||  
QY 6136 GTGAACAAGGGAAGAGGGCA 6155  
Db |||||||

RESULT 9  
AX503780  
LOCUS AX503780 6189 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 10 from Patent WO0226826.  
ACCESSION AX503780  
VERSION AX503780.1 GI:23385965  
KEYWORDS Homo sapiens (human)  
SOURCE

ORGANISM Homo sapiens  
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Gerlach, V.L., Macdougall, J.R., Smithson, G., Millet, I., Stone, D., Gunther, E., Ellerman, K., Grosse, W.M., Alsobrook, J.P., Lepley, D.M., Burgess, C.E., Padigaru, M., Kekuda, R., Spytek, K.A., Leach, M.D. and Shimkets, R.A.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0226826-A 10 04-APR-2002;  
FEATURES Curagen Corporation (US)  
source location/Qualifiers  
1. 6189  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 85.7%; Score 5631; DB 6; Length 6189;  
Best Local Similarity 95.3%; Pred. No. 0;  
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;  
QY 19 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 78  
Db |||||||  
QY 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 60  
Db |||||||  
QY 79 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCAA 138  
Db |||||||  
QY 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCAA 120  
Db |||||||  
QY 139 CAGCAGATGCTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTTTTGA 198  
Db |||||||  
QY 121 CAGCAGATGCTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTTTTGA 180  
Db |||||||  
QY 199 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGAAGTGT 258  
Db |||||||  
QY 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGAAGTGT 240  
Db |||||||  
QY 259 TCCGACACCATAGCTAGTTTACAGGAGCTCCAGCCTTCGGCAAAGGACTTCGAAGTCAGA 318  
Db |||||||  
QY 241 TCCGACACCATAGCTAGTTTACAGGAGCTCCAGCCTTCGGCAAAGGACTTCGAAGTCAGA 300  
Db |||||||  
QY 319 AGTCTGTAGTGTGCTGCTCTTTGCTGAAGTGCAGGTGCTTAAGAGAGAAAGCAACGGG 378  
Db |||||||  
QY 301 AGTCTGTAGTGTGCTGCTCTTTGCTGAAGTGCAGGTGCTTAAGAGAGAAAGCAACGGG 360  
Db |||||||  
QY 379 GACATCTATGCTGAAGTGAAGAGAGAGGCTTTATTGGCCAGGAGGAGGTTTCA 438  
Db |||||||  
QY 361 GACATCTATGCTGAAGTGAAGAGAGAGGCTTTATTGGCCAGGAGGAGGTTTCA 420  
Db |||||||  
QY 439 TTTTGTGAGAGAGCGGGAACATATTATCTCGAAGCAACAGCCCGTGGATCCCCCAATTA 498  
Db |||||||  
QY 421 TTTTGTGAGAGAGCGGGAACATATTATCTCGAAGCAACAGCCCGTGGATCCCCCAATTA 480  
Db |||||||  
QY 499 CAGTATGCTTTCAGGACAAATAATCACCTTTATCTGATGAGAGAAATATCAGCCTGGAGG 558  
Db |||||||  
QY 481 CAGTATGCTTTCAGGACAAATAATCACCTTTATCTGATGAGAGAAATATCAGCCTGGAGG 540  
Db |||||||  
QY 559 GACTTGTGCTCACCTTTTGAATAGATATGAGACCCAGTTAGATGAAACCTGATACAGTTT 618  
Db |||||||  
QY 541 GACTTGTGCTCACCTTTTGAATAGATATGAGACCCAGTTAGATGAAACCTGATACAGTTT 600  
Db |||||||  
QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCACTGATGGGATACGTCATCGA 678  
Db |||||||  
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCACTGATGGGATACGTCATCGG 660  
Db |||||||  
QY 679 GACATCAGCTGAGACATTTCTGTTGACCGCAGGACACATCAAGCTGGTGGATTTT 738  
Db |||||||  
QY 661 GACATCAGCTGAGACATTTCTGTTGACCGCAGGACACATCAAGCTGGTGGATTTT 720  
Db |||||||  
QY 739 GGATCTGCGCGGAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTTGGACC 798  
Db |||||||  
QY 721 GGATCTGCGCGGAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTTGGACC 777  
Db |||||||

QY 799 CCAGATTACATGGCTCTCGAAGTGTGACTGTGATGAACCGGGATGGAAGGACCTTAC 858  
Db 778 CCAGATTACATGGCTCTCGAAGTGTGACTGTGATGAACCGGGATGGAAGGACCTTAC 837  
QY 859 GGCCTGGACTGTGACTGGTGGTCACTGGGCGGTGATGCTATGAGATGATTTATGGGAGA 918  
Db 838 GGCCTGGACTGTGACTGGTGGTCACTGGGCGGTGATGCTATGAGATGATTTATGGGAGA 897  
QY 919 TCCCGCTTCGACAGAGGGAACCTTCGCCAGAACCTTCAATAACATPATGAATTCACGGG 978  
Db 898 TCCCGCTTCGACAGAGGGAACCTTCGCCAGAACCTTCAATAACATPATGAATTCACGGG 957  
QY 979 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTTCCTTGATCTGATTCARAGC 1038  
Db 958 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTTCCTTGATCTGATTCARAGC 1017  
QY 1039 TTGTTGTGGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGTGCGCATCTTTC 1098  
Db 1018 TTGTTGTGGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGTGCGCATCTTTC 1077  
QY 1099 TCTAAATTTGACTGGAACAAANTTCGTAACCTCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1158  
Db 1078 TCTAAATTTGACTGGAACAAANTTCGTAACCTCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1137  
QY 1159 TCTGACGATGACACCTCCAAATTTTGATGAACCCAGAGAGAAATTCGTGGGTTCATCTCT 1218  
Db 1138 TCTGACGATGACACCTCCAAATTTTGATGAACCCAGAGAGAAATTCGTGGGTTCATCTCT 1197  
QY 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTCG 1278  
Db 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTCG 1257  
QY 1279 TACAGCAAGGCACTGGGATCTTCGTAGATCTGAGTCTGTTGTGTCGGGTCTGACTCC 1338  
Db 1258 TACAGCAAGGCACTGGGATCTTCGTAGATCTGAGTCTGTTGTGTCGGGTCTGACTCC 1317  
QY 1339 CTTGCCAAGACTAGCTCCATGGAAGAACTTCTCATCAAGCAAAAGACTCAAGAC 1398  
Db 1318 CTTGCCAAGACTAGCTCCATGGAAGAACTTCTCATCAAGCAAAAGACTCAAGAC 1377  
QY 1399 TCTCAGGCAAGTGTCAAGATGAGCAGGAAATGACCCGTTACATCGAGAGTGCA 1458  
Db 1378 TCTCAGGCAAGTGTCAAGATGAGCAGGAAATGACCCGTTACATCGAGAGTGCA 1437  
QY 1459 GAGGTGGAGGCTGTCTTAGTCAGAGGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1518  
Db 1438 GAGGTGGAGGCTGTCTTAGTCAGAGGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1497  
QY 1519 TCCCTCTTGAGCAGGACCTTGCTACCTACATCAGAAATGCGATGAGTTAAAGGAAAT 1578  
Db 1498 TCCCTCTTGAGCAGGACCTTGCTACCTACATCAGAAATGCGATGAGTTAAAGGAAAT 1557  
QY 1579 TTGGAGCAAGCAGGATGGAGTGTCCAGAGGAGTGAAGGACTCAGCTTCTCCAT 1638  
Db 1558 TTGGAGCAAGCAGGATGGAGTGTCCAGAGGAGTGAAGGACTCAGCTTCTCCAT 1617  
QY 1639 GATATCAGAGCAGAGCCGGAAGTCCAGAGAAATCAAGAGCAGGAGTACAGGCTCAA 1698  
Db 1618 GATATCAGAGCAGAGCCGGAAGTCCAGAGAAATCAAGAGCAGGAGTACAGGCTCAA 1677  
QY 1699 GTGGAAGAAATCAGGTTCATGATGAATCAGTTGGAGAGGATCTTGCTCAGCAGAGAGA 1758  
Db 1678 GTGGAAGAAATCAGGTTCATGATGAATCAGTTGGAGAGGATCTTGCTCAGCAGAGAGA 1737  
QY 1759 CGGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTAAGAAATTCAG 1818  
Db 1738 CGGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTAAGAAATTCAG 1797  
QY 1819 CGGAAAGCAGAGATGTACAGATTAACCTGTTGAAGGCTTAGGATCAAGGAGGCTGAA 1878  
Db 1798 CGGAAAGCAGAGATGTACAGATTAACCTGTTGAAGGCTTAGGATCAAGGAGGCTGAA 1857

QY 1879 GTGGGAGAAATATCGGAAACTGGAGAAAGATCAATGCTTGAGCAGAGCTCAAAATTCAGGAG 1938  
Db 1858 GTGGGAGAAATATCGGAAACTGGAGAAAGATCAATGCTTGAGCAGAGCTCAAAATTCAGGAG 1917  
QY 1939 CTCCAAGAGAAACTGGAGAAAGGCTG----- 1963  
Db 1918 CTCCAAGAGAAACTGGAGAAAGGCTGTTAAAGCCAGACCGAGGCCACCGAGCTGCTGCAG 1977  
QY 1964 -----CAAGAGCGAGCCGAGAGGGAGCTGGAGAAAGCTGCGAAGCCGAGAG 2010  
Db 1978 AATATCGCCAGGCAAGAGGAGCGAGCGAGAGGGAGCTGGAGAAAGCTGCGAAGCCGAGAG 2037  
QY 2011 GATTCTTCTCAAGGCAATCAGAAAGAGTGGTGGAAAGCTGAGGAAAGCGCCCAATTCCTCG 2070  
Db 2038 GATTCTTCTCAAGGCAATCAGAAAGAGTGGTGGAAAGCTGAGGAAAGCGCCCAATTCCTCG 2097  
QY 2071 GAGAAACAAGTAAAGAGACTTAGAGACCATGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2130  
Db 2098 GAGAAACAAGTAAAGAGACTTAGAGACCATGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2157  
QY 2131 ATCCAGACAAATCCCAACAGATCCAGCCACACTAGAAGTGCACCTGAAACAGAAAGAG 2190  
Db 2158 ATCCAGACAAATCCCAACAGATCCAGCCACACTAGAAGTGCACCTGAAACAGAAAGAG 2217  
QY 2191 AAACATCGGAGGCCCAAGTCTCAGCCACACTAGAAGTGCACCTGAAACAGAAAGAG 2250  
Db 2218 AAACATCGGAGGCCCAAGTCTCAGCCACACTAGAAGTGCACCTGAAACAGAAAGAG 2277  
QY 2251 CAGCACTATGAGAAAGATTTAAAGTTGTGGCAATCAGATTAAGAAAGACCTTGGCTGAC 2310  
Db 2278 CAGCACTATGAGAAAGATTTAAAGTATTGGCAATCAGATTAAGAAAGACCTTGGCTGAC 2337  
QY 2311 AAGGAGACACTGAGAAACATGATCGAGAGACAGAGAGGAGGCCCATGAGAAAGGGCAAA 2370  
Db 2338 AAGGAGACACTGAGAAACATGATCGAGAGACAGAGAGGAGGCCCATGAGAAAGGGCAAA 2397  
QY 2371 ATTCTCAGCAACAGAAAGCGATGATCAATGCTATGATTTCCAAGATCAGATCCCTGGAA 2430  
Db 2398 ATTCTCAGCAACAGAAAGCGATGATCAATGCTATGATTTCCAAGATCAGATCCCTGGAA 2457  
QY 2431 CAGAGGATTTGGAACCTGCTGAAGCCCAATAAACTTGCAGCAATAGCAGTCTTTTACC 2490  
Db 2458 CAGAGGATTTGGAACCTGCTGAAGCCCAATAAACTTGCAGCAATAGCAGTCTTTTACC 2517  
QY 2491 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550  
Db 2518 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577  
QY 2551 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAATACTGAGGAGCAGCTGGAG 2610  
Db 2578 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAATACTGAGGAGCAGCTGGAG 2637  
QY 2611 AAGATCAGCCCAACAGACAGTGAACAAGATTCGGCTGCTGGAACCTGGAGACAAGATTG 2670  
Db 2638 AAGATCAGCCCAACAGACAGTGAACAAGATTCGGCTGCTGGAACCTGGAGACAAGATTG 2697  
QY 2671 CGGAGGTGAGTCTAGAGACAGAGAGCAGAACTCGAGCTTAAGCGCAGCTCAGAG 2730  
Db 2698 CGGAGGTGAGTCTAGAGACAGAGAGCAGAACTCGAGCTTAAGCGCAGCTCAGAG 2757  
QY 2731 CTACAGCTCTCCCTGAGGAGCGGAGTCACTGTTCAAGCCCTGCAAGCTGCACGGCG 2790  
Db 2758 CTACAGCTCTCCCTGAGGAGCGGAGTCACTGTTCAAGCCCTGCAAGCTGCACGGCG 2817  
QY 2791 GCCTCTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTTGAAGAGACCAACAGAGAGCT 2850  
Db 2818 GCCTCTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTTGAAGAGACCAACAGAGAGCT 2877  
QY 2851 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAATTCAGCGCAAAATTTGATGCT 2910  
Db 2878 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAATTCAGCGCAAAATTTGATGCT 2937  
QY 2911 CTTTCGTAACAGCTGTACTGTAAATCAGAGACCTGGAGGAGCAGCTAAACCAAGCTGACCGAG 2970



4981	QY	CCCAACAATTTTGGAGCTGTCTCAAGGGTGCGCCACTTTGTTTGGGGCAGGCAAGATTGAGAAC	5040
5113	Db	CCCAACAATTTTGAAGCTGTCTCAAGGGTGCGCCACTTTGTTTGGGGCAGGCAAGATTGAGAAC	5172
5041	QY	GGGCTCTGCATCTGTGCAGGCATGCCAGCAAAAGTCTGTCATTCTCCGCTACAAACGAAAAC	5100
5173	Db	GGGCTCTGCATCTGTGCAGGCATGCCAGCAAAAGTCTGTCATTCTCCGCTACAAACGAAAAC	5232
5101	QY	CTCAGCAAAATCTGCATCTCCGAAAGAGATAGAGACCTTCAGAGCCCTGCAGCTGTATCCAC	5160
5233	Db	CTCAGCAAAATCTGCATCTCCGAAAGAGATAGAGACCTTCAGAGCCCTGCAGCTGTATCCAC	5292
5161	QY	TTCAACCAATTCACAGTATCCCTCATTTGGAACCAATAAATTTTACGAAATTCGACATGAGCAG	5220
5293	Db	TTCAACCAATTCACAGTATCCCTCATTTGGAACCAATAAATTTTACGAAATTCGACATGAGCAG	5352
5221	QY	TACACGCTCAGGAATTCCTTGGATAAGAAATGACCAATTCCTTGGCACCTGCTGTGTTGCC	5280
5353	Db	TACACGCTCAGGAATTCCTTGGATAAGAAATGACCAATTCCTTGGCACCTGCTGTGTTGCC	5412
5281	QY	GCCTCTTCCAAACAGCTTCCTCTGCTCTCAATCGTGCAGGTGAAACGCGCAGGGCAGCAGAG	5340
5413	Db	GCCTCTTCCAAACAGCTTCCTCTGCTCTCAATCGTGCAGGTGAAACGCGCAGGGCAGCAGAG	5472
5341	QY	GAGTACTTGTCTGTTTCCACGAATTTGGAGTGTTCGTGGAATTCCTTACGGAAGACGTAGC	5400
5473	Db	GAGTACTTGTCTGTTTCCACGAATTTGGAGTGTTCGTGGAATTCCTTACGGAAGACGTAGC	5532
5401	QY	CGCACAGACATCTCAAGTGGAGTCGCTTACCTTTTGGCCCTTTCGCTACAGAGAACCTTAT	5460
5533	Db	CGCACAGACATCTCAAGTGGAGTCGCTTACCTTTTGGCCCTTTCGCTACAGAGAACCTTAT	5592
5461	QY	CTGTTTGTGACCCACTTCAACTCAGTCGAAGTAATTCAGATCCAGCAGCAGCTCCTCAGCA	5520
5593	Db	CTGTTTGTGACCCACTTCAACTCAGTCGAAGTAATTCAGATCCAGCAGCAGCTCCTCAGCA	5652
5521	QY	GGGACCCCTGCCCGAGCGTACCTGGACATCCCGAAACCCGCGCTACCTTGGGCCCTGCCATT	5580
5653	Db	GGGACCCCTGCCCGAGCGTACCTGGACATCCCGAAACCCGCGCTACCTTGGGCCCTGCCATT	5712
5581	QY	TCCTCAGGACGATTTACTTTGGCGTCCTCATACAGGATAAATTAAGGTTCATTTCGTGC	5640
5713	Db	TCCTCAGGACGATTTACTTTGGCGTCCTCATACAGGATAAATTAAGGTTCATTTCGTGC	5772
5641	QY	AAGGGAAACCTCGTGAAGAGTCGGGACATGAACACACCGGGGGCCGCTCCACCTCCCGC	5700
5773	Db	AAGGGAAACCTCGTGAAGAGTCGGGACATGAACACACCGGGGGCCGCTCCACCTCCCGC	5832
5701	QY	AGCAGCCCCAAAGCGAGGCCACCCACCGTAACACGAGCACATCAACAGCCGCTGGCC	5760
5833	Db	AGCAGCCCCAAAGCGAGGCCACCCACCGTAACACGAGCACATCAACAGCCGCTGGCC	5892
5761	QY	TCCAGCCACGCGCCGCGGAAGCCCGCAGCCACCGCGAGAGCCAAAGCACCCACCCG	5820
5893	Db	TCCAGCCACGCGCCGCGGAAGCCCGCAGCCACCGCGAGAGCCAAAGCACCCACCCG	5952
5821	QY	TACCGAGGGCGGACCGAGCTCGCGAGGGACAAAGTCTCTCTGGCCGCCCTTCGAGCGA	5880
5953	Db	TACCGAGGGCGGACCGAGCTCGCGAGGGACAAAGTCTCTCTGGCCGCCCTTCGAGCGA	6012
5881	QY	GAGAGTCCCCCGGCCGATGCTCAGCACGCGGAGAGAGCGGTCCCCCGGGAGGCTGTTT	5940
6013	Db	GAGAGTCCCCCGGCCGATGCTCAGCACGCGGAGAGAGCGGTCCCCCGGGAGGCTGTTT	6072
5941	QY	GAAGACAGCAGCAGGGGCGGCTGCGTGGCGGAGCCGTGAGGACCCCGCTGTCCCAAGTG	6000
6073	Db	GAAGACAGCAGCAGGGGCGGCTGCGTGGCGGAGCCGTGAGGACCCCGCTGTCCCAAGTG	6132
6001	QY	AACAAGGAAGAGGCGAGATGCTCTCAAGTTTTCACGGTTTAAACTGTCACT	6055
6133	Db	AACAAGGTGTGGACCCAGTCTTTCAGTATAAATCTCAGCCAGAAAAAACCACTCCT	6187

[illegible]

Db 601 TACTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTTCATCTCTGATGGGATACCTGTCATCGG 660  
QY 679 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 738  
Db 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720  
QY 739 GATCTGCGCGGAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTTGGGACC 798  
Db 721 GATCTGCGCGGAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTTGGGACC 777  
QY 799 CAGATTTACATGCTCTCTGAAGTGTGACTGTGATCAACCGGGATGGAAAAGCCACTAC 858  
Db 778 CAGATTTACATGCTCTCTGAAGTGTGACTGTGATCAACCGGGATGGAAAAGCCACTAC 837  
QY 859 GGCTTGAATGTGACTGGTGGTCAAGTGGGCGTGAATTCCTATGAGATGAATTTATGGGAGA 918  
Db 838 GGCTTGAATGTGACTGGTGGTCAAGTGGGCGTGAATTCCTATGAGATGAATTTATGGGAGA 897  
QY 919 TCCGCTTTCGACAGGGAACCTCTGCGCAGAACTTCAATAAATATGAATTTCCAGCGG 978  
Db 898 TCCGCTTTCGACAGGGAACCTCTGCGCAGAACTTCAATAAATATGAATTTCCAGCGG 957  
QY 979 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1038  
Db 958 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1017  
QY 1039 TTGTTGCGGCGCAGAAAGAGACTCAAGTTTGAAGTGTGCTGCCATCCTTCTTC 1098  
Db 1018 TTGTTGCGGCGCAGAAAGAGACTCAAGTTTGAAGTGTGCTGCCATCCTTCTTC 1077  
QY 1099 TCTAAAATTTGACTGGAACCAATCTCTCTCCGCTTCTGTTCCCACTCAAG 1158  
Db 1078 TCTAAAATTTGACTGGAACCAATCTCTCTCCGCTTCTGTTCCCACTCAAG 1137  
QY 1159 TCTGAGATGACACTTCCAAATTTGATGAACACAGAGAATTCGTGGGTTTCATCCTCT 1218  
Db 1138 TCTGAGATGACACTTCCAAATTTGATGAACACAGAGAATTCGTGGGTTTCATCCTCT 1197  
QY 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGTTTGTGGGTTTCG 1278  
Db 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGTTTGTGGGTTTCG 1257  
QY 1279 TACAGCAAGCCTGCGGATCTTGTGATGATCTGCTGTTGTGTCGGTTCGACTCC 1338  
Db 1258 TACAGCAAGCCTGCGGATCTTGTGATGATCTGCTGTTGTGTCGGTTCGACTCC 1317  
QY 1339 CCGTGCAGCTGAGCTTCCATGGAAGAACTTCTCATCAAGCAAGAGCTACAAGAC 1398  
Db 1318 CCGTGCAGCTGAGCTTCCATGGAAGAACTTCTCATCAAGCAAGAGCTACAAGAC 1377  
QY 1399 TCTCAGCAAGTGTCAAGATGAGCAGGAATGACCCGGTTACATCGGAGGTGCA 1458  
Db 1378 TCTCAGCAAGTGTCAAGATGAGCAGGAATGACCCGGTTACATCGGAGGTGCA 1437  
QY 1459 GAGGTGGAGCTGTGCTTAGTCAGAGGAGGTGGAGCTCAAGCCCTCAGACTCAGAGA 1518  
Db 1438 GAGGTGGAGCTGTGCTTAGTCAGAGGAGGTGGAGCTCAAGCCCTCAGACTCAGAGA 1497  
QY 1519 TCCCTCTGAGCAGGACCTTGTACTCATCATCAGAAATGCAAGTGTAAAGCGAAGT 1578  
Db 1498 TCCCTCTGAGCAGGACCTTGTACTCATCATCAGAAATGCAAGTGTAAAGCGAAGT 1557  
QY 1579 TTGGAGCAAGCAGGATGAGGTGTCCAGAGGAGTGAACAAGCACTGAGCTTCTCCAT 1638  
Db 1558 TTGGAGCAAGCAGGATGAGGTGTCCAGAGGAGTGAACAAGCACTGAGCTTCTCCAT 1617  
QY 1639 GATATCAGAGCAGAGCCGGAAGCTTCCAGAAATCAAAAGCAGGAGTACCAGGCTCAA 1698  
Db 1618 GATATCAGAGCAGAGCCGGAAGCTTCCAGAAATCAAAAGCAGGAGTACCAGGCTCAA 1677  
QY 1699 GTGGAAGAAATGAGTTGATGAATTCAGTTGGAAAGAGATCTTGTCTCAGCAAGAGA 1758  
Db 1678 GTGGAAGAAATGAGTTGATGAATTCAGTTGGAAAGAGATCTTGTCTCAGCAAGAGA 1737

QY 1759 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAATTCAG 1818  
Db 1738 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAATTCAG 1797  
QY 1819 CGGAACCGCAGAGATCTCAGCATAACTCTTGAAGCTAAGGATCAAGGGAACCTCGAA 1878  
Db 1798 CGGAACCGCAGAGATCTCAGCATAACTCTTGAAGCTAAGGATCAAGGGAACCTCGAA 1857  
QY 1879 GTGGAGAATATCGAACTCGGAAGAATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938  
Db 1858 GTGGAGAATATCGAACTCGGAAGAATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1917  
QY 1939 CTCCAGAGAACTGGAGAGGCT----- 1962  
Db 1918 CTCCAGAGAACTGGAGAAAGGCTGTAAGCCAGCAGCAGGAGGCCACCGAGCTGCTGCAG 1977  
QY 1963 -----GCAAGGAGCGAGCGGAGGAGCTGGAGAAGCTGCAGAAACCGAGAG 2010  
Db 1978 AATATCCGCGCAGCAAGGAGCGAGCGGAGGAGCTGGAGAAGCTGCAGAAACCGAGAG 2037  
QY 2011 GATTCCTTCTGAGGCATCAAAAGAGAGCTGGTGAAGCTGAGGAACGCCGCTTCTCTG 2070  
Db 2038 GATTCCTTCTGAGGCATCAAAAGAGAGCTGGTGAAGCTGAGGAACGCCGCTTCTCTG 2097  
QY 2071 GAGAAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGACTGAAGAGTAC 2130  
Db 2098 GAGAAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGACTGAAGAGTAC 2157  
QY 2131 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGAGCTCGAAGAG 2190  
Db 2158 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGAGCTCGAAGAG 2217  
QY 2191 AAACATCGGAGGCGCAAGTCTCAGCCAGCACCCTAGAGTGCACCTGAAACAGAAAGAG 2250  
Db 2218 AAACATCGGAGGCGCAAGTCTCAGCCAGCACCCTAGAGTGCACCTGAAACAGAAAGAG 2277  
QY 2251 CAGCACTATGAGGAAAGATTTAAAGTGTTCGAACAATCAGATAAAGAAAGACCTGGGTGAC 2310  
Db 2278 CAGCACTATGAGGAAAGATTTAAAGTGTTCGAACAATCAGATAAAGAAAGACCTGGGTGAC 2337  
QY 2311 AAGAGACACTGAGNAACATGATCAGCCAGCACCCTAGAGAGGAGGCCCATGAGAGGCAAA 2370  
Db 2338 AAGAGACACTGAGNAACATGATCAGAGACACAGGAGGAGGCCCATGAGAGGCAAA 2397  
QY 2371 ATTCTCAGCGAAACAGAAAGGCGGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2430  
Db 2398 ATTCTCAGCGAAACAGAAAGGCGGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2457  
QY 2431 CAGAGGATTTGGAACTGTCTGAAGCCAATAAATTGACAGCAATAGCAGTCTTTTACC 2490  
Db 2458 CAGAGGATTTGGAACTGTCTGAAGCCAATAAATTGACAGCAATAGCAGTCTTTTACC 2517  
QY 2491 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTAC 2550  
Db 2518 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTAC 2577  
QY 2551 CTGAGACACAGGCTGGGAAATTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG 2610  
Db 2578 CTGAGACACAGGCTGGGAAATTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG 2637  
QY 2611 AAGATCAGCCACCAAGACCAAGTACAGAAATCGGCTGCTGGAATCTGGAGACAAGATTG 2670  
Db 2638 AAGATCAGCCACCAAGACCAAGTACAGAAATCGGCTGCTGGAATCTGGAGACAAGATTG 2697  
QY 2671 CGGAGGTGAGTCTAGAGCAGGAGCAGAAACTGGAGCTCAAGCGCAGCTCAGCAGAG 2730  
Db 2698 CGGAGGTGAGTCTAGAGCAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCAGCAGAG 2757  
QY 2731 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTCAGAGCTTGCAGGCTGCACGGGCG 2790  
Db 2758 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTCAGAGCTTGCAGGCTGCACGGGCG 2817

QY 2791 GCCCTGGAGCCAGCTTCGCCAGCGGAGACAGAGCTGGAGAGACCACAGCAAGCT 2850  
Db 2818 GCCCTGGAGCCAGCTTCGCCAGCGGAGACAGAGCTGGAGAGACCACAGCAAGCT 2877  
QY 2851 GAAGAGGAGATCCAGGCACTACCGGACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2910  
Db 2878 GAAGAGGAGATCCAGGCACTACCGGACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2937  
QY 2911 CTTCTAACAGTGTACTGTANTACAGACCTGGAGGAGCAGCTTAAACCAGGTGACCGAG 2970  
Db 2938 CTTCTAACAGTGTACTGTANTACAGACCTGGAGGAGCAGCTTAAACCAGGTGACCGAG 2997  
QY 2971 GACAAACGCTGAACCTCAACCAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3030  
Db 2998 GACAAACGCTGAACCTCAACCAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3057  
QY 3031 GCGCCAAACGACGAGATGTGTAACAATGCGAAGTGAAGTGGACCACTCTCCGCGGAGATC 3090  
Db 3058 GCGCCAAACGACGAGATGTGTAACAATGCGAAGTGAAGTGGACCACTCTCCGCGGAGATC 3117  
QY 3091 ACGGAACGAGAGATCGAGCTTACCAGCCAGAGCAACAGATGGAGGCTCTGAAGACCAAG 3150  
Db 3118 ACGGAACGAGAGATCGAGCTTACCAGCCAGAGCAACAGATGGAGGCTCTGAAGHCCAG 3177  
QY 3151 TGCACCATGCTGGAGGAAACAGGTCTATGGATTGGAGGCCCTAAACGATGAGCTGCTAGAA 3210  
Db 3178 TGCACCATGCTGGAGGAAACAGGTCTATGGATTGGAGGCCCTAAACGATGAGCTGCTAGAA 3237  
QY 3211 AAAGAGCGGAGTGGAGGCCCTGGAGGAGCGTCTCGGCTGATGAGAAATCCAGCTTTCAG 3270  
Db 3238 AAAGAGCGGAGTGGAGGCCCTGGAGGAGCGTCTCGGCTGATGAGAAATCCAGCTTTCAG 3297  
QY 3271 TGTCTGGGTTCCGAGAGCTGCAGAGAAATGTGTGACACCGAGAAACAGAGAGGCGCAGAGCC 3330  
Db 3298 TGTCTGGGTTCCGAGAGCTGCAGAGGATGTGTGACACCGAGAAACAGAGAGGCGCAGAGCC 3357  
QY 3331 GATCAGCGGATCACCAGGCTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGGACACAAGGCT 3390  
Db 3358 GATCAGCGGATCACCAGGCTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGGACACAAGGCT 3417  
QY 3391 GAGATTCTCGCTCTCAGCAGGCTCTCAAAGAGCAGAAAGCTGAAGCGCAGAGGCTCTCT 3450  
Db 3418 GAGATTCTCGCTCTCAGCAGGCTCTCAAAGAGCAGAAAGCTGAATGCTTGAATGAATGCCGAAGCTTA 3537  
QY 3451 GACAAAGCTCAATGACTGGAGAAAGCATGCTATGCTTGAATGAATGCCGAAGCTTA 3510  
Db 3478 GACAAAGCTCAATGACTGGAGAAAGCATGCTATGCTTGAATGAATGCCGAAGCTTA 3537  
QY 3511 CAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGCCAAA 3570  
Db 3538 CAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGCCAAA 3597  
QY 3571 TTACAGCAGCATGGAGCTGCAGAAAAATCACATTTTCCGCTCTGACTCAAGGACTGCAA 3630  
Db 3598 TTACAGCAGCATGGAGCTGCAGAAAAATCACATTTTCCGCTCTGACTCAAGGACTGCAA 3657  
QY 3631 GAAGCTTAGATCGGGCTGATCTACTGAAGCAGAAAAAGAGTGAATTCGAGTATCAGCTG 3690  
Db 3658 GAAGCTTAGATCGGGCTGATCTACTGAAGCAGAAAAAGAGTGAATTCGAGTATCAGCTG 3717  
QY 3691 GAAACATTTCAGGTTCTCTATTCTCATGAAAAGGTGAAAAATGGAAGGCACTATTTCCTCAA 3750  
Db 3718 GAAACATTTCAGGTTCTCTATTCTCATGAAAAGGTGAAAAATGGAAGGCACTATTTCCTCAA 3777  
QY 3751 CAAAACAACTCATTTGATTTTCGAAAGCAAAATGGACCAACCTGCTAAAAAGAAAAAG 3810  
Db 3778 CAAAACAACTCATTTGATTTTCGAAAGCAAAATGGACCAACCTGCTAAAAAGAAAAAG - - 3835  
QY 3811 GGTTTATTAGTCGAGGAAAGAGGACCTGCTTTACCCACACAGGTTCTCTGAGTAC 3870  
Db 3836 -----AGTGCCTCTGAGTAC 3852  
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGTGTGCAGAGTGAAGAGCC 3930

Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3912  
QY 3931 CTTTCAGAAAGACCCGCATTCGAGCTTCGGTTCGCCCGGAGGAAAGCTGCCACCGCAAGCA 3990  
Db 3913 CTTTCAGAAAGACCCGCATTCGAGCTTCGGTTCGCCCGGAGGAAAGCTGCCACCGCAAGCA 3972  
QY 3991 ACGGACCAACCAACCAATCCAGCCAGCACCGCGAGGAGCAGATCGCCATGTCCGCC 4050  
Db 3973 ACGGACCAACCAACCAATCCAGCCAGCACCGCGAGGAGCAGATCGCCATGTCTGCC 4032  
QY 4051 ATCGTGGGTCGCCAGAGACACCGCCAGTGCATGAGCCTGTGCGCCCGCCCATCCAGC 4110  
Db 4033 ATCGTGGGTCGCCAGAGCACCGCCAGTGCATGAGCCTGTGCGCCCGCCCATCCAGC 4092  
QY 4111 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTTCGGCGTCTTAAGGAAACGATGCAC 4170  
Db 4093 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTTCGGCGTCTTAAGGAAACGATGCAC 4152  
QY 4171 CACAATATTCTCAACCGATTCAACGTAGGACTGAAATGCGAGGCCACAAAGTGTGCTGTG 4230  
Db 4153 CACAATATTCTCAACCGATTCAACGTAGGACTGAAATGCGAGGCCACAAAGTGTGCTGTG 4212  
QY 4231 TGTCTGATACCGTGCATCTTTGGACGCGAGGCATCCAAATGTCTCGAAATGTCAAGTGTATG 4290  
Db 4213 TGTCTGATACCGTGCATCTTTGGACGCGAGGCATCCAAATGTCTAGAAATGTCAAGTGTATG 4272  
QY 4291 TGTCAACCCCAAGTGTCTCAGTCTTGCAGCCAGCCTGCGGCTTGCCTGTGAATATGCC 4350  
Db 4273 TGTCAACCCCAAGTGTCTCAGTCTTGCAGCCAGCCTGCGGCTTGCCTGTGAATATGCC 4332  
QY 4351 ACACACTTCAACCGAGGCTTCTCCGCTGACAAATGAACCTCCCGAGTCTCCAGACCAAG 4410  
Db 4333 ACACACTTCAACCGAGGCTTCTCCGCTGACAAATGAACCTCCCGAGTCTCCAGACCAAG 4392  
QY 4411 GAGCCAGCAGAGCTTGCACCTGGAGGGTGGATGAAGTGGCCAGGAATACAAACGA 4470  
Db 4393 GAGCCAGCAGAGCTTGCACCTGGAGGGTGGATGAAGTGGCCAGGAATACAAACGA 4452  
QY 4471 GGACAGCAAGGCTGGGACAGGAAAGTACATTTCTCTGGAGGGAATCAAAAGTCTCATTTAT 4530  
Db 4453 GGACAGCAAGGCTGGGACAGGAAAGTACATTTCTCTGGAGGGAATCAAAAGTCTCATTTAT 4512  
QY 4531 GACAAATGAAGCCAGAGAGCTGGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4590  
Db 4513 GACAAATGAAGCCAGAGAGCTGGACAGAGCCCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4572  
QY 4591 GACGGGATCTATCTATTCTATGCTGCGTGGTGGTCTTCCGAACTCGCAAAATACAGCCAAA 4650  
Db 4573 GACGGGATCTATCTATTCTATGCTGCGTGGTGGTCTTCCGAACTCGCAAAATACAGCCAAA 4632  
QY 4651 GCA----- 4653  
Db 4633 GCAGATGCCCATACATACCTGAAGATGGAATCTCACCCGCACACCACTGCTGCGCCCGGG 4692  
QY 4654 ----- 4653  
Db 4693 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCTTCA 4752  
QY 4654 -----GAAAAAGCAGAAAGCTGATAAATCG 4680  
Db 4753 GAATCAGTGTGCGAGGTTGGAGAGTTTCTAGGAAAAAAGCAGAGCTGATGATAAATCG 4812  
QY 4681 CTTGGAAAACTCCCTGTGTAACCTGGAAGTGTAGCCGTCTAGACATGAATGCACGCTG 4740  
Db 4813 CTTGGAAAACTCCCTGTGTAACCTGGAAGTGTAGCCGTCTAGACATGAATGCACGCTG 4872  
QY 4741 CCCTTCAGTACACAGGTGGTGTGGTGGGACCGAGGAAGGGCTCTACGCCCTGAATGTC 4800  
Db 4873 CCCTTCAGTACACAGGTGTGGTGGGACCGAGGAAGGGCTCTACGCCCTGAATGTC 4932  
QY 4801 TTGAAAACTCCCTTAACCCCATGTCCAGGAATTTGGAGCAGCTTCTCCAAATTTATATTC 4860

4933 TTGAATACTCCCTAACCCATGTCCTCCAGGATGTTGGAGCAGTCTTCCAAATTTATATTATC 4992  
4861 AAGGACCTGAGAGCTACTCATGATAGCAGAGAGAGAGCGGCGCATGTGTCTTGTGGAC 4920  
4993 AAGGACCTGAGAGCTACTCATGATAGCAGAGTGAAGAGCGGCGCATGTGTCTTGTGGAC 5052  
4921 GTGAAGAAAGTGAAGAGCTCCCTGGCGCCAGTCCACCTGCTGCGCCAGCCGACATCTCA 4980  
5053 GTGAAGAAAGTGAAGAGCTCCCTGGCGCCAGTCCACCTGCTGCGCCAGCCGACATCTCA 5112  
4981 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5040  
5113 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5172  
5041 GGGCTCTGCACTCTGTGAGGCGATGCCAGCAAGTCTGATCTTCCGCTACACAGCAAAAC 5100  
5173 GGGCTCTGCACTCTGTGAGGCGATGCCAGCAAGTCTGATCTTCCGCTACACAGCAAAAC 5232  
5101 CTCAGCAATATCTCCATCTCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5160  
5233 CTCAGCAATATCTCCATCTCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5292  
5161 TTCACCAATATACAGTATCTCATCTGGAACCAATAAATTTCTACGAATCGACATGAAGCAG 5220  
5293 TTCACCAATATACAGTATCTCATCTGGAACCAATAAATTTCTACGAATCGACATGAAGCAG 5352  
5221 TACACGCTCAGGAATCTCCGATAAGATGACCATCTTCCCTGGCACCTGCTGTGTTGCC 5280  
5353 TACACGCTCAGGAATCTCCGATAAGATGACCATCTTCCCTGGCACCTGCTGTGTTGCC 5412  
5281 GCCTTTTCCAAAGCTTCCCTGCTCAATCTGTCAGGTGAACAGCGCAGGCGAGGAG 5340  
5413 GCCTTTTCCAAAGCTTCCCTGCTCAATCTGTCAGGTGAACAGCGCAGGCGAGGAG 5472  
5341 GAGTACTTGTGTTTCCAGCAATTTGGAGTCTTCTGATTTCTTACGGAAGAGCTAGC 5400  
5473 GAGTACTTGTGTTTCCAGCAATTTGGAGTCTTCTGATTTCTTACGGAAGAGCTAGC 5532  
5401 CGACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT 5460  
5533 CGACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT 5592  
5461 CTGTTTGTGACCCACTTCACTCACTGGAAGTAAATGAGATCCAGCAGCTCCTCAGCA 5520  
5593 CTGTTTGTGACCCACTTCACTCACTGGAAGTAAATGAGATCCAGCAGCTCCTCAGCA 5652  
5521 GGGACCCCTGCCGAGCGTACCTGAGCATCTCCGAAACCCCGGCTACCTGGGCGCCCTGCCATT 5580  
5653 GGGACCCCTGCCGAGCGTACCTGAGCATCTCCGAAACCCCGGCTACCTGGGCGCCCTGCCATT 5712  
5581 TCCTCAGGACGATTTACTTGGGCTCCTCATACAGGATAAATTAAGGTCATTTGCTGC 5640  
5713 TCCTCAGGACGATTTACTTGGGCTCCTCATACAGGATAAATTAAGGTCATTTGCTGC 5772  
5641 AAGGGAAACCTGCTGAAGAGTCCGCGACTGAAACACACCGGGCGCCCTCCACCTCCCGC 5700  
5773 AAGGGAAACCTGCTGAAGAGTCCGCGACTGAAACACACCGGGCGCCCTCCACCTCCCGC 5832  
5701 AGACGCCCCAACAGCGAGGCGCCACCGCATCAACAGGACATCAACAGCGCGGTGGCC 5760  
5833 AGACGCCCCAACAGCGAGGCGCCACCGCATCAACAGGACATCAACAGCGCGGTGGCC 5892  
5761 TCACGCGCCAGCGCGCCGAGGCGCCACCGCATCAACAGGAGCGCAAGCAGACCCACCGC 5820  
5893 TCACGCGCCAGCGCGCCGAGGCGCCACCGCATCAACAGGAGCGCAAGCAGACCCACCGC 5952  
5821 TACCGGAGGGCGGACCGAGTCTGCGAGGACAGTCTCTTGGCGCGCCCTCGAGGCGA 5880  
5953 TACCGGAGGGCGGACCGAGTCTGCGAGGACAGTCTCTTGGCGCGCCCTCGAGGCGA 6012  
5881 GAGAGTCTCCCGCGGATGCTCAGACGCGGAGAGCGGTCCCCCGGAGGCTGTTT 5940  
6013 GAGAGTCTCCCGCGGATGCTCAGACGCGGAGAGCGGTCCCCCGGAGGCTGTTT 6072

QY 5941 GAAGACAGCAGCGCGCGCTGCCTGCGGGAGCCCTGAGGACCCCGCTGTCCCGAGTG 6000  
Db 6073 GAAGACAGCAGCAGCGCGCGCTGCCTGCGGGAGCCCTGAGGACCCCGCTGTCCCGAGTG 6132  
QY 6001 AACAGGGAAGAGGCGCA 6017  
Db 6133 AACAGGTCAGGCGAGCA 6149  
RESULT 11  
LOCUS AY257469 8576 bp mRNA linear PRI 23-APR-2003  
DEFINITION Homo sapiens rho/rac-interacting citron kinase (CIT) mRNA, complete cds.  
ACCESSION AY257469  
VERSION AY257469.1 GI:30088969  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 8576)  
AUTHORS Huang,C.Q., Wu,S.L., Shan,Y.X., Liu,S. and Xiao,P.J.  
TITLE Direct Submission  
JOURNAL Submitted (18-MAR-2003) Department of Biochemistry and Molecular Biology, Basic Medicine, Suzhou University, Renming Road 48, Suzhou, Jiangsu 215007, China  
FEATURES  
source 1..8576  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q24.23"  
1..8576  
/gene="CIT"  
/notes="synonym: CR1K"  
54..6137  
/gene="Citr"  
/notes="serine/threonine protein kinases"  
/codon\_start=1  
/product="rho/rac-interacting citron kinase"  
/protein\_id="AA013528.1"  
/db\_xref="GI:30088970"  
/translation="MLKFKYQARNPLDAGAAEPIASASRLNLFPOGKPPFTQQQMS  
PLREGILDALFVLPFECSQALMKIKHVSFNFKYSDTIAELOELQPSAKQPEVRS  
VGGHFAEVQVREKATGDIYAMKMKKALLAQEQVFFEBERNILSRSTSPWIPQL  
OYAFQDNHLYLVMEYQPGGDLNRYEDLDENLIQYLAELIYLVHVSVMGCV  
HRDIPENIIVDRTHGKLVDFGSAAMNSNMVNAKLPIGTPDYNAPEVLTVMGDD  
KTYGLDCDMSVGVIAEMIVGSPSPAGTSARTFNNIMNFORLKFDPDPKVSDF  
LDLIQLCCQKRLKAFEGLCCHPFSKIDWNINRSPFPFVTLSDDDTSNDFSE  
KNWSVSPQSPSPSGEELPFVGFVSYSKALGILGRSESVVSGSDSPAKTSMEKK  
LLIKSELQSDQKCHKMEQMTLRRHVSVEAVLSQKEVELKASETSQRLLEDLA  
TYTESLSLRSLQARMESQDDKALQLLHDIRQSKLQIKQEQYQAQVEML  
MNQLEEDIVARRRDLVSELSRSLAAEFKFKATECOHKLKAKQKQPEVEY  
AKKLEINAEQOLKIQLQELKLVKASTEELLQNIQAKERAERELEKLNQVDS  
SEGIRKLVAESELEKHEAEOVAOHLEVLKQEQVVEEIKVLNDQIKKLAQKE  
TLENMQRHEEAHEKQKILSBOKAMINAMDSKIRSEQLIVELSEANKLANSSIFT  
QRNKAQEMISELRQOKFYLETQAGLEAQRNKRLEQEKSHQSHQDNRLLELET  
RLREVSLEEQKLEQLQTLQSLQESQTLQALQARAALQSOLQKQNTFELSET  
TAAAEIIEALTHREIQRKFDALRNCTVIITDEEQLNQLTDEADNLAELNQLYLSK  
QLDEAGANDEIVQLRSEVDHLREITEREMQTSQTEALKTCTMLEEQMDLE  
ALNDELLEKPEWAEWSVLGDEKQFECRVELQMLDTEKQSRARQORITESQV  
VELAVKEKAEILALQALKEOKLKAESLSDKLNLEKHAMLENARSILQOKLETER  
ELQRLLEQAKLQOQMDLQKNIHIFLTQGLQELADRLDKLTERSDLYQLENIQVL  
YSEKQVMEGTITSQTKLIDLOAKMDQAKKKKGLFSRKEDPALPTQVLPQYQNLK  
LALKEKARCAELEALQKTRIELRSEEAHRKATMDHPHSTPATASQQAIAIV  
RSEHPQFSAMSLAPSSRRKESSTPEFSRRKERMHNIHPRFVGNLMNRKCAV  
CLDTVHFGRQAKCLQVMCHPKCTCLPATCGLPAEYATHTFEACRDKMNSPGLQ  
TKPSSSLHLEGMKYPNRNKGQGDWKYIVLEGSKVLIVDNEAREAGQPEVEFE  
LCIPDGDSVTHGAVGASELANAKADVPYILKMESHPHTTCWPGRTLYLLAPSPDKQ  
RWTALSVGVRGVRERKAEADAKLLGNLSLLKLEGGDRDMMNCTLPFSDQVVLVGT



QY 1981 AGGAGCTGGAAGCTGCAAGACCGAGAGGATTCCTTCTGAAGGCATCAAGAAAGAGCTG 2040  
Db 2015 --GCCACCAGCTGTGCAAGATATCCGCCAGGCAAGAGGAGCGAGCCGAGAGGAGCTG 2072  
QY 2041 GTGGAGCTGAGGAGCGCCCATCTCTGGGAAACAAGGTAAAGAGATAGAGACCATG 2100  
Db 2073 GAGAAGCTGAGAACCGAGAGGATTCCTTCTGAAGGCATCAGAAAGAGCTGTGG----- 2127  
QY 2101 GAGCGTAGAAGAAACAGACTGAAGGATGACATCCAGACAAAATCCCAACAGATCCAGCAG 2160  
Db 2128 ----- 2127  
QY 2161 ATGGCTGATAAATCTGGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGCCGAG 2220  
Db 2128 -----AAGCTGAGGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGCCGAG 2177  
QY 2221 CAGCTAGAGCTGACCTGAAACAGAAAGAGCAGCACTATGAGGAAAAGATTAAAGTGTG 2280  
Db 2178 CACCTAGAGCTGACCTGAAACAGAAAGAGCAGCACTATGAGGAAAAGATTAAAGTGTG 2237  
QY 2281 GACAATCAGATAAAGAAAGCCTGGCTGACAAGAGACACTGGAGAAACATGATGAGAGA 2340  
Db 2238 GACAATCAGATAAAGAAAGCCTGGCTGACAAGAGACACTGGAGAAACATGATGAGAGA 2297  
QY 2341 CAGCAGGAGGAGCCCATGAGAGAGGCAAAATTTCTCAGCGAAACAGAAAGCGATGATCAAT 2400  
Db 2298 CAGCAGGAGGAGGCCCATGAGAGAGGCAAAATTTCTCAGCGAAACAGAAAGCGATGATCAAT 2357  
QY 2401 GCTATGATTCAGAGTCCAGATCCCTGGAAACAGAGAGATTGTGGAATGTCTGAAGCCAAAT 2460  
Db 2358 GCTATGATTCAGAGTCCAGATCCCTGGAAACAGAGAGATTGTGGAATGTCTGAAGCCAAAT 2417  
QY 2461 AACTTGCACCAATAGCAGCTTTTACCCAAAGGAACATGAAGCCCAAGAGAGATG 2520  
Db 2418 AACTTGCACCAATAGCAGCTTTTACCCAAAGGAACATGAAGCCCAAGAGAGATG 2477  
QY 2521 ATTCTGAACTCAGGCAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCC 2580  
Db 2478 ATTCTGAACTCAGGCAACAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCC 2537  
QY 2581 CAGAACCGAAACTGGAGGAGCTGGAGAGATCAGCCACAGACACACAGCTGACAG 2640  
Db 2538 CAGAACCGAAACTGGAGGAGCTGGAGAGATCAGCCACAGACACACAGCTGACAG 2597  
QY 2641 AATCGGCTGTGGAACCTGGAGACAAGATTGCGGGAGGTGAGTCTAGAGCACGAGAGCAG 2700  
Db 2598 AATCGGCTGTGGAACCTGGAGACAAGATTGCGGGAGGTGAGTCTAGAGCACGAGAGCAG 2657  
QY 2701 AACTGGAGCTCAAGCGCCAGCTCAGAGCTACAGCTCTCCTCTGAGAGCGCGAGTCA 2760  
Db 2658 AACTGGAGCTCAAGCGCCAGCTCAGAGCTACAGCTCTCCTCTGAGAGCGCGAGTCA 2717  
QY 2761 CAGTTGACAGCTGAGAGCTGACAGGCGGCTGGAGAGCAGCTTCGCCAGGCGAAG 2820  
Db 2718 CAGTTGACAGCTGAGAGCTGACAGGCGGCTGGAGAGCAGCTTCGCCAGGCGAAG 2777  
QY 2821 ACAGAGCTGGAAGAGACCAAGAGAGCTGGAAGAGGATCCAGCACTCAGCGACAT 2880  
Db 2778 ACAGAGCTGGAAGAGACCAAGAGAGCTGGAAGAGGAGATCCAGCACTCAGCGACAT 2837  
QY 2881 AGAGATGAATCCAGCGCAATTTGATGCTCTTCGTAACAGCTGTAATCAAGAG 2940  
Db 2838 AGAGATGAATCCAGCGCAATTTGATGCTCTTCGTAACAGCTGTAATCAAGAG 2897  
QY 2941 CTGGAGGAGCAGTAAACAGCTGACCGAGGACACCGCTGAATCAACAAACAAACTTC 3000  
Db 2898 CTGGAGGAGCAGTAAACAGCTGACCGAGGACACACCGCTGAATCAACAAACAAACTTC 2957  
QY 3001 TACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCCCAACGAGAGATTGTAACCTGCGA 3060  
Db 2958 TACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCCCAACGAGAGATTGTAACCTGCGA 3017  
QY 3061 AGTGAAGTGGAACCATCTCCGCCGGAGATCAAGGAAACAGAGATGAGCTTACCAGCCAG 3120

Db 3018 AGTGAAGTGGAACCATCTCCGCCGGAGATCAGGAAACGAGAGATGACAGCTTACCAGCCAG 3077  
QY 3121 AAGCAAAACGATGGAGGCTCTGAAGACCAACGTCACACCTGTGGAGGAACAGGTCATGGAT 3180  
Db 3078 AAGCAAAACGATGGAGGCTCTGAAGACCAACGTCACACCTGTGGAGGAACAGGTCATGGAT 3137  
QY 3181 TTGAGGCGCTTAAACGATGAGCTCTGAAGAAAGAGCGGAGCTGGAGGCTGGAGAGC 3240  
Db 3138 TTGAGGCGCTTAAACGATGAGCTCTGAAGAAAGAGCGGAGCTGGAGGCTGGAGAGC 3197  
QY 3241 GTCCTGGGTGATGAGAAATCCAGTTTGTGCTGGGTTCCAGAGCTGCAGAGATCCTG 3300  
Db 3198 GTCCTGGGTGATGAGAAATCCAGTTTGTGCTGGGTTCCAGAGCTGCAGAGATGCTG 3257  
QY 3301 GACACCGAGAAACAGAGCAGGGGAGAGCCGATCAGCGGATCACCGAGTCTCCGCCAGGTG 3360  
Db 3258 GACACCGAGAAACAGAGCAGGGGAGAGCCGATCAGCGGATCACCGAGTCTCCGCCAGGTG 3317  
QY 3361 GTGAGCTGGCAGTGAAGGAGCAAGGCTGAGATTCTCGCTCTGCAGAGGCTCTCATA 3420  
Db 3318 GTGAGCTGGCAGTGAAGGAGCAAGGCTGAGATTCTCGCTCTGCAGAGGCTCTCATA 3377  
QY 3421 GAGCAGAGCTGAAGGCGGAGAGCCTCTCTGACAGCTCAATGACCTGGAGAAAGACAT 3480  
Db 3378 GAGCAGAGCTGAAGGCGGAGAGCCTCTCTGACAGCTCAATGACCTGGAGAAAGACAT 3437  
QY 3481 GCTATGCTGAAATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGAACGAGCTC 3540  
Db 3438 GCTATGCTGAAATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGAACGAGCTC 3497  
QY 3541 AAAACAGAGGCTTCTGGAAGAGCAAGCCAAATTCAGCAGCAGATGGACCTGCAGAAAAAT 3600  
Db 3498 AAAACAGAGGCTTCTGGAAGAGCAAGCCAAATTCAGCAGCAGATGGACCTGCAGAAAAAT 3557  
QY 3601 CACATTTTCCGTCTGACTCAAGAGCTCAAGAGACTAGATCGGCTGATCTACTGAAG 3660  
Db 3558 CACATTTTCCGTCTGACTCAAGAGCTCAAGAGACTAGATCGGCTGATCTACTGAAG 3617  
QY 3661 ACAGAAAGAGTGAATTTGGAGTATCAGCTGGAACAACTTCAGGTTCTCTATTTCTCATGAA 3720  
Db 3618 ACAGAAAGAGTGAATTTGGAGTATCAGCTGGAACAACTTCAGGTTCTCTATTTCTCATGAA 3677  
QY 3721 AAGTGAAAAATGGAAGGCACTATTTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGCC 3780  
Db 3678 AAGTGAAAAATGGAAGGCACTATTTCTCAACAAACCAAACTCATTTGATTTTCTGCAAGCC 3737  
QY 3781 AAAATGGACCAACCTGCTAAAAAGAAAAGGTTATTTAGTTCGACGGAAGAGGACCT 3840  
Db 3738 AAAATGGACCAACCTGCTAAAAAGAAAAGGTTATTTAGTTCGACGGAAGAGGACCT 3797  
QY 3841 GCTTTACCCACACAGGTTCTCTGAGTACAATGAGTGAAGCTGGCCCTGGAGAGGAG 3900  
Db 3798 GCTTTACCCACACAGGTTCTCTGAGTACAATGAGTGAAGCTGGCCCTGGAGAGGAG 3857  
QY 3901 AAAGCTCGCTGTCAGAGCTAGAGGAGCCCTTCAGNAGACCCGATCGAGCTCCGCTCC 3960  
Db 3858 AAAGCTCGCTGTCAGAGCTAGAGGAGCCCTTCAGNAGACCCGATCGAGCTCCGCTCC 3917  
QY 3961 GCCCGGAGGAAGCTGCCACCCGCAAGCAACCGGACCCACCAACCCATCCAGCCAGCC 4020  
Db 3918 GCCCGGAGGAAGCTGCCACCCGCAAGCAACCGGACCCACCAACCCATCCAGCCAGCC 3977  
QY 4021 ACCCGAGGAGCAGAGATCGCCATGTCGCCATGTCGGGTCCGAGAGCAGCAGCCAGT 4080  
Db 3978 ACCCGAGGAGCAGAGATCGCCATGTCGCCATGTCGGGTCCGAGAGCAGCAGCCAGT 4037  
QY 4081 GCCATGAGCTGCTGGGCCCGCCATCCAGCGCAGAAAGAGTCTTCAACTCCAGAGGAA 4140  
Db 4038 GCCATGAGCTGCTGGGCCCGCCATCCAGCGCAGAAAGAGTCTTCAACTCCAGAGGAA 4097  
QY 4141 TTTAGTGGCGCTTTAAGGAACCGATGCAACATTTCTCACCAGATTCAACGTAGGA 4200





Db	1861	GTGGAGATATGCGAAATCGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1920
Qy	1939	CTCCAGAGAAACTGGAGAGGCT-----	1962
Db	1921	CTCCAGAGAAACTGGAGAGGCTGTAAAGCCAGCAGCGAGGCCACCGAGTGCTGCAG	1980
Qy	1963	-----GCAAAGAGGAGCGCGAGCGGAGGAGCTGGAGAACTCGAGAACCTCGAGAACCGAGAG	2010
Db	1981	AATATCCGCGCAGCAAGAGGAGCGAGCGGAGGAGCTGGAGAACTCGAGAACCGAGAG	2040
Qy	2011	GATTTCTTGAAGGCATCAGAAAGAGCTGTGGAGCTGAGGAAACCGCGCATTTCTTG	2070
Db	2041	GATTTCTTGAAGGCATCAGAAAGAGCTGTGGAGCTGAGGAAACCGCGCATTTCTTG	2100
Qy	2071	GAGAACAGGTAAAGAGACTAGAGCCATGGAGCCTAGAGAAACAGACTGAGAGATGAC	2130
Db	2101	GAGAACAGGTAAAGAGACTAGAGCCATGGAGCCTAGAGAAACAGACTGAGAGATGAC	2160
Qy	2131	ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAG	2190
Db	2161	ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAG	2220
Qy	2191	AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG	2250
Db	2221	AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG	2280
Qy	2251	CAGCACTATGAGGAAAGATTAAAGTTGTGACAATCAGATAAAGAAAGACTGGCTGAC	2310
Db	2281	CAGCACTATGAGGAAAGATTAAAGTTGTGACAATCAGATAAAGAAAGACTGGCTGAC	2340
Qy	2311	AAGGAGACACTGGAGAACATGATGAGAGACACGAGGAGGAGGCCATGAGAGGGCAAA	2370
Db	2341	AAGGAGACACTGGAGAACATGATGAGAGACACGAGGAGGAGGCCATGAGAGGGCAAA	2400
Qy	2371	ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGGATTCACAGATCAGATCCCTGGAA	2430
Db	2401	ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGGATTCACAGATCAGATCCCTGGAA	2460
Qy	2431	CAGAGGATTGTGGAACTCTCTGAAGCCCAATAACTTGAGCAAAATAGCAGTCTTTTACC	2490
Db	2461	CAGAGGATTGTGGAACTCTCTGAAGCCCAATAACTTGAGCAAAATAGCAGTCTTTTACC	2520
Qy	2491	CAAAGGACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC	2550
Db	2521	CAAAGGACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC	2580
Qy	2551	CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG	2610
Db	2581	CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG	2640
Qy	2611	AAGATCAGCCACCAAGACACAGTGAAGAATCGGCTGTGGAACTGGAGACAGATTTG	2670
Db	2641	AAGATCAGCCACCAAGACACAGTGAAGAATCGGCTGTGGAACTGGAGACAGATTTG	2700
Qy	2671	CGGAGGTCAGTCTAGAGCACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAGAG	2730
Db	2701	CGGAGGTCAGTCTAGAGCACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAGAG	2760
Qy	2731	CTACAGCTCTCCCTCAGAGCGCAGTCAAGTTGACAGCCCTGACGGCTGCACGGGCG	2790
Db	2761	CTACAGCTCTCCCTCAGAGCGCAGTCAAGTTGACAGCCCTGACGGCTGCACGGGCG	2820
Qy	2791	GCCTTGGAGAGCCAGCTTCGCAGCGAGACAGAGCTGGAGAGACCACAGCAGAGCT	2850
Db	2821	GCCTTGGAGAGCCAGCTTCGCAGCGAGAGACAGAGCTGGAGAGACCACAGCAGAGCT	2880
Qy	2851	GAAGAGGAGATCCAGGCACTCAGGCAATAGAGATGAAATCCAGCGCAAAATTTGATGCT	2910
Db	2881	GAAGAGGAGATCCAGGCACTCAGGCAATAGAGATGAAATCCAGCGCAAAATTTGATGCT	2940
Qy	2911	CTTCGTAAACAGTGTACTGTAAATCAGACCTGGAGGAGAGCTTAAACAGCTGACCGAG	2970
Db	2941	CTTCGTAAACAGTGTACTGTAAATCAGACCTGGAGGAGAGCTTAAACAGCTGACCGAG	3000
Qy	2971	GACAACTGAACTCAACAAACCAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCT	3030
Db	3001	GACAACTGAACTCAACAAACCAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCT	3060
Qy	3031	GGCCCAACACAGAGATTGTACAACTCGAAAGTGAAGTGACCATCTCCCGCGGAGATC	3090
Db	3061	GGCCCAACACAGAGATTGTACAACTCGAAAGTGAAGTGACCATCTCCCGCGGAGATC	3120
Qy	3091	ACGGAACGAGAGATGACAGCTTACGAGCCAGAAACAAACGATGAGGCTCTGAAAGACCA	3150
Db	3121	ACGGAACGAGAGATGACAGCTTACGAGCCAGAAACAAACGATGAGGCTCTGAAAGACCA	3180
Qy	3151	TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA	3210
Db	3181	TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA	3240
Qy	3211	AAAGAGCGGAGTGGGAGGCTTGGAGAGCGTCTCGGTGATGAGAAATCCCGATTGAG	3270
Db	3241	AAAGAGCGGAGTGGGAGGCTTGGAGAGCGTCTCGGTGATGAGAAATCCCGATTGAG	3300
Qy	3271	TGTCGGGTTTCGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCAGGGCGAGGCC	3330
Db	3301	TGTCGGGTTTCGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCAGGGCGAGGCC	3360
Qy	3331	GATCAGCGGATCACCGAGTCTCGCCAGTGGTGGAGCTGGCAGTGAAGGAGCACAAGGCT	3390
Db	3361	GATCAGCGGATCACCGAGTCTCGCCAGTGGTGGAGCTGGCAGTGAAGGAGCACAAGGCT	3420
Qy	3391	GAGATTTCTCGCTCTGAGCAGGCTCTCAAAGACAGAGCTGAAGCCGAGAGCGCTTCT	3450
Db	3421	GAGATTTCTCGCTCTGAGCAGGCTCTCAAAGACAGAGCTGAAGCCGAGAGCGCTTCT	3480
Qy	3451	GACAAGCTCAATGACCTGGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3510
Db	3481	GACAAGCTCAATGACCTGGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3540
Qy	3511	CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGGCAAA	3570
Db	3541	CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGGCAAA	3600
Qy	3571	TTACAGCAGCAGATGACCTGCGAGAAATTCATATTTCCGCTGCTGCTGCTGCTGCTGCTG	3630
Db	3601	TTACAGCAGCAGATGACCTGCGAGAAATTCATATTTCCGCTGCTGCTGCTGCTGCTGCTG	3660
Qy	3631	GAAGCTTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTGAAGTGAAGTGAAGTGAAGT	3690
Db	3661	GAAGCTTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTGAAGTGAAGTGAAGTGAAGT	3720
Qy	3691	GAAAAATTCAGGTTCTCTATTCTCATGAAAAGTGAATAATGGAAGGCACTATTTCTCAA	3750
Db	3721	GAAAAATTCAGGTTCTCTATTCTCATGAAAAGTGAATAATGGAAGGCACTATTTCTCAA	3780
Qy	3751	CAAAACCAAACTGATTTCTGCAAGCCAAATGGACCACTCTGCTGCTGCTGCTGCTGCTGCTG	3810
Db	3781	CAAAACCAAACTGATTTCTGCAAGCCAAATGGACCACTCTGCTGCTGCTGCTGCTGCTGCTG	3838
Qy	3811	GGTTTATTTAGTCGACGGAAAGAGACCTGCTTTTACCACACAGGTTCTCTGCAAGTAC	3870
Db	3839	-----AGGTTCTCTGCAAGTAC	3855
Qy	3871	AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3930
Db	3856	AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3915
Qy	3931	CTTCAGAAACCCGATCGAGCTTCGGTCCGCGCGGAGGAAAGCTGCCCAACCGCAAGCA	3990
Db	3916	CTTCAGAAACCCGATCGAGCTTCGGTCCGCGCGGAGGAAAGCTGCCCAACCGCAAGCA	3975
Qy	3991	ACGGAACCCACACCCATCCAGCCAGCCACCGCAGGAGCAGATCGCCATGTCGCGC	4050
Db	3976	ACGGAACCCACACCCATCCAGCCAGCCACCGCAGGAGCAGATCGCCATGTCGCGC	4035

```
QY 4051 ATGTCGCGTCCGACAGACACAGCCAGTGCATGAGCCTGCTGGCCCGCCATCCAGC 4110
Db 4036 ATGTCGCGTCCGACAGACACAGCCAGTGCATGAGCCTGCTGGCCCGCCATCCAGC 4095
QY 4111 CGCAGAAAGGAGTCTCAACTCCAGAGGAATTTAGTCGCGGCTTTAAGGAACGCATGCAC 4170
Db 4096 CGCAGAAAGGAGTCTCAACTCCAGAGGAATTTAGTCGCGGCTTTAAGGAACGCATGCAC 4155
QY 4171 CACAATATTTCCTCACCGAATTCACAGTAGGACTGAACATCGAGGCCACAAAGTGTCTGTG 4230
Db 4156 CACAATATTTCCTCACCGAATTCACAGTAGGACTGAACATCGAGGCCACAAAGTGTCTGTG 4215
QY 4231 TGTCTGATACCGTGACATTTGGACCCAGGAGATCCAAATGCTCGAATGTCAGTGATG 4290
Db 4216 TGTCTGATACCGTGACATTTGGACCCAGGAGATCCAAATGCTCGAATGTCAGTGATG 4275
QY 4291 TGTCAACCCAAAGTCTCCAGTCTTGCAGGCCACCTGGCGGCTTGCTGCTGAATATGCC 4350
Db 4276 TGTCAACCCAAAGTCTCCAGTCTTGCAGGCCACCTGGCGGCTTGCTGCTGAATATGCC 4335
QY 4351 ACACATTCACAGAGGCTTTCGCTGTGACAAAATGAATCTCCAGAGTCTCCAGACCAAG 4410
Db 4336 ACACACTTCACAGAGGCTTTCGCTGTGACAAAATGAATCTCCAGAGTCTCCAGACCAAG 4395
QY 4411 GAGCCAGAGAGGCTTTCACCTGGAGGCTGATGAAGTGGCCAGGAAATACAAACGA 4470
Db 4396 GAGCCAGAGAGGCTTTCACCTGGAGGCTGATGAAGTGGCCAGGAAATACAAACGA 4455
QY 4471 GGAAGCAAGGCTGGACAGGAGTACATTTGCTGGAGGATCAAAAGTCTCATTTAT 4530
Db 4456 GGAAGCAAGGCTGGACAGGAGTACATTTGCTGGAGGATCAAAAGTCTCATTTAT 4515
QY 4531 GACAAATGAGCCAGAGAGCTGGACAGAGCGCGTGGAGAAATTCAGCTGCGCTTCCC 4590
Db 4516 GACAAATGAGCCAGAGAGCTGGACAGAGCGCGTGGAGAAATTCAGCTGCGCTTCCC 4575
QY 4591 GAGCGGAGTATCTATTTCATGTCGCGTTGCTTCGAACTCGCAATACAGCCAAA 4650
Db 4576 GAGCGGAGTATCTATTTCATGTCGCGTTGCTTCGAACTCGCAATACAGCCAAA 4635
QY 4651 GGA----- 4653
Db 4636 GCAGATGTCCCATACATACATGGAATCTCACCCGACACACACCTGCTGGCCCGGG 4695
QY 4654 ----- 4653
Db 4696 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTACCCGCTTA 4755
QY 4654 -----GAAAAAGCAGAGCTGATGCTGCTAAACTG 4680
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTCTTAGGGAAAAAGCAGAGCTGATGCTAAACTG 4815
QY 4681 CTTTGGAACTCCCTGCTGAACTGGAGGTGATGACCGCTAGACATGAATGACGCTG 4740
Db 4816 CTTTGGAACTCCCTGCTGAACTGGAGGTGATGACCGCTAGACATGAATGACGCTG 4875
QY 4741 CCCTTCAGTGACAGGTGTTGTTGGTGGCCACCGAGGAAGGCTCTACGCCCTGAATGTC 4800
Db 4876 CCCTTCAGTGACAGGTGTTGTTGGTGGCCACCGAGGAAGGCTCTACGCCCTGAATGTC 4935
QY 4801 TTGAAAACCTCCCTAACCCATGTCACAGGAATGGAGAGTCTTCCAAATTTATATTATC 4860
Db 4936 TTGAAAACCTCCCTAACCCATGTCACAGGAATGGAGAGTCTTCCAAATTTATATTATC 4995
QY 4861 AAGGACCTGGAGAGCTACTCATATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 4920
Db 4996 AAGGACCTGGAGAGCTACTCATATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 5055
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4980
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 5115
QY 4981 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGGCAGGCAAGATTGAGAAC 5040
```

```
Db 5116 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5175
QY 5041 GGGCTCTGTCATCTGTGTCAGCCATGCCAGCAAAAGTGTGTATTTCCGCTACAAAGAAAC 5100
Db 5176 GGGCTCTGTCATCTGTGTCAGCCATGCCAGCAAAAGTGTGTATTTCCGCTACAAAGAAAC 5235
QY 5101 CTAGAGAAATACATGATCCCGAAAGATAGACACCTCGAGCCCTCGAGCTGTATCCAC 5160
Db 5236 CTAGAGAAATACATGATCCCGAAAGATAGACACCTCGAGCCCTCGAGCTGTATCCAC 5295
QY 5161 TTCAACAAATACAGTATTCCTCATTTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5220
Db 5296 TTCAACAAATACAGTATTCCTCATTTGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5355
QY 5221 TACAGCTCGAGGAATTCCTGATTAAGATGACATTCCTTGGCACCTGCTGTGTTGCC 5280
Db 5356 TACAGCTCGAGGAATTCCTGATTAAGATGACATTCCTTGGCACCTGCTGTGTTGCC 5415
QY 5281 GCCTTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGCAGGCGACGAGAG 5340
Db 5416 GCCTTTCCAAACAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGCAGGCGACGAGAG 5475
QY 5341 GAGTACTGTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5400
Db 5476 GAGTACTGTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5535
QY 5401 CGCACAGACCATCTCAAGTGGAGTTCCTTACCTTTGGCCTTTCCCTACAGAGAACCTAT 5460
Db 5536 CGCACAGACCATCTCAAGTGGAGTTCCTTACCTTTGGCCTTTCCCTACAGAGAACCTAT 5595
QY 5461 CTGTTTGTGACCCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCAAGCTCCTCAGCA 5520
Db 5596 CTGTTTGTGACCCCACTTCAACTCACTCGAAGTAAATGAGATCCAGGCAAGCTCCTCAGCA 5655
QY 5521 GGGACCCCTCCCGAGGCTACCTGGACATCCCAACCCGCGCTACCTGGGCGCTGCCATT 5580
Db 5656 GGGACCCCTCCCGAGGCTACCTGGACATCCCAACCCGCGCTACCTGGGCGCTGCCATT 5715
QY 5581 TCCTCAGAGCGATTTACTTTGGGCTCTCATACAGGATAAAATTAAGGTCATTGCTGC 5640
Db 5716 TCCTCAGAGCGATTTACTTTGGGCTCTCATACAGGATAAAATTAAGGTCATTGCTGC 5775
QY 5641 AAGGAAACCTCTGTGAAGAGTCCGGACATGAACACCCGCGGCGCTCCACCTCCCGC 5700
Db 5776 AAGGAAACCTCTGTGAAGAGTCCGGACATGAACACCCGCGGCGCTCCACCTCCCGC 5835
QY 5701 AG 5702
Db 5836 AG 5837
```

```
RESULT 13
AF086824 6954 bp mRNA linear ROD 11-NOV-1998
LOCUS Mus musculus rho/rac-interacting citron kinase (Crik) mRNA,
DEFINITION complete cds.
ACCESSION AF086824
VERSION AF086824.1 GI:3599508
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 6954)
AUTHORS Di Cunto,F., Calautti,E., Hsiao,J., Ong,L., Topley,G., Turco,E. and
Dotto,G.P.
TITLE Citron rho-interacting kinase, a novel tissue-specific ser/thr
kinase encompassing the rho-Rac-binding protein Citron
J. Biol. Chem. 273 (45), 29706-29711 (1998)
JOURNAL 99009084
MEDLINE PUBMED 9792683
REFERENCE 2 (bases 1 to 6954)
```

AUTHORS Di Cunto,F., Calautti,E., Hsiao,J., Ong,L., Topley,G., Turco,E. and Dotto,G.P.  
TITLE Direct Submission  
JOURNAL Submitted (25-AUG-1998) Biology, University of Torino, Via Santena 5 bis, Torino, TO 10126, Italy  
FEATURES Location/Qualifiers  
source 1..6954  
/organism="Mus musculus"  
/mol\_type="rRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="spleen"  
1..6954  
/gene="Crik"  
772..6939  
/gene="Crik"  
/function="protein kinase"  
/note="CR1K; rho/rac binding protein; alternatively spliced form of citron encoded by GenBank Accession Number AF039218; similar to Mus musculus citron-K encoded by GenBank Accession Number AF070066; similar to Rattus norvegicus citron-K encoded by GenBank Accession Number AF070065"  
/codon\_start=1  
/product="rho/rac-interacting citron kinase"  
/protein\_id="AAC7823.1"  
/db\_xref="GI:3599509"  
/translation="MLKFKYGVNPPPEASBPASRASRLNLPFGKPPMLTQOOMS  
ALSRGMLDALPALPECSOPALMKMKVHSFVQKYSOTIAELRELQASDFVRSL  
VQGFHFAEVVVEKATGDVYAKMKKALLAQEVAFPEERNILSRSTPMIPOL  
YQAFQDKNLYLWMEYQPGDQLSLNRYEDQIDSMIQFYLAELLLAVHSHVNGVY  
HADIKPNLLIIRDTLIDFVGSAAKNSNKKVDKFLIGFIDVWAPVFLVIMEDRR  
GYDGLCDWMSVGVAYEMVYKTFTEGTSARTFNINMNFQFLKFPDDPKVSESL  
DILQSLCKQKELAFGLCCHFFPARTDMNNIRNPPFPVFLKSDSDTFNDFPEK  
NSWAFILCPKPLAFSGELPFVGFYSKALGYLGRSESVSSLDSPAKVSSMEKIL  
LIKSELQDSQKCHKMEQEMTRHRRYSEVAVLQKEVELKASSETORSILLEQAT  
VITECSLRSLEQARMVYSOEDDKALQLLHDIREQSRKLQEIKEQVQAOVEERLM  
MNQBEDLYSARRSDLEIYSELRESRLAAEFKRYANSCQHKLMKAKQCGPEVGEYS  
KLEKINAEQOLKIQLQELKELKAVKASTATELLQNIQAKERABERLEKLNHRDSS  
EGIKKLVAEERERLSEKNVRLTMEERENLDDIQKSEIQQWADKILLEEK  
HREAQVASHLEVLHKKQEHYEEKI KVLNDQIKDLADKESLENMQRHEEAEHK  
KILSEQAMINADNRKLEORIVELSEANKLAANSLSFTORNMQAEMISELRQK  
KYLETQACKLAQNRKLEOLEKISHQSDSKRLLELELREVSLEHEBQKLELK  
ROLTELQSLQRESQLTALAQARALESQLEQAKTELEHETAEAEESIQAHTARD  
IQRFDALNRSTCIDLEQNLQEDNENLNQNFYLSKOLDASGANBEIVLRS  
EYDHLREITREMQLTSQKTEALKTCTLMEQVLDLEALNDELLEKQEWAWR  
SVLGDEQSEFCRQELQMLTEKQSRADQRTESQVVELAVKHEKABILALQ  
ALKEQKLAESLDKLNLEKKHLENNARSLOQKLEKRELKORLEEEQAKLQOM  
DLQKNIHFLITQLOALDRADLLTERSDYLEQENIOVLYSHKVKMEGTISOQTK  
LIDFLQAKVDQPAKKKVPLOYNELKLALEKAKCAELEALOKTRIELRSAREEA  
HRKATHPHSPATARQIAMSIVRSEHQPASMSLLAPSSRKSSSTPEEPSRR  
LKERMHNLPFRNGLNWRATKAVCLDTVHFGQASKLEECQVNCBPXSCICLPAT  
CGLPBYATHTEAFQDNQNSPGQSKEPSSHLHGEMKVPNNKRGQGWDERYI  
VLEGSVLIYDNEAREAGQRPVEEFELCLPDGDSVTHGAVGASELANAKADVPYILK  
MESHPHTTCWPGRTLYLLAPSPDKQRTVALNSVAVGVRSEKAEADAKLLGNSLL  
KLEGGDRDLWNCITLPESDQVLVGTBEGLYALNVLNLSUTHPGIGAVFOIYIHKDLE  
KLMIAGERALCLVDVKVKOSLAOSHLPADPVSNTFEAVKSHLPFAAGKIENSL  
CICAMPSPKVIILYRNDNLSKVCIRKEIETSPESCIFHTNYSILGINKFYEIDMKQ  
YIUDEFLKNDHSLAPAFASSNSFPFVIVANSAGQREYLLCFHFPFVDSYGR  
RSTDDLKWSRPLFAFAYREPLFVTHFNSLVEIEQARSLSGSPARYLEIPNRYL  
GPAISSGAIYLASQYQKLVICQKGNLVKESQHQHVRPSTSRSPNKRGPPTNEH  
ITKRVASSPAPPGSPHSPSTPHRYRDRGRTLELRDKSPGRPLEREKSPGRMLST  
RRRSRPGRLFEDSSRGLPAGAVRTPLSQNVKWDQSSV"

817 GCCTCCGAGCCCAATTCAGTCGGGCTCCAGGCTAAATCTCTTCTCCAGGGGAAACCG 876  
124 CCCTTTATGACTCAACAGCAGATGTCTCTCTTTCCGAGAGAGGATATTAGATGCCTC 183  
877 CCCCTCATGACTCAACAGCAGATGTCTCTCTTTCCGAGAGAGGATGCTAGAGCCCTC 936  
184 TTTCTTCTCTTTGAAGAATGCAGTCAGCTCAGCTCTCTGATGAAGATTAAAGCAGCTGAGCAAC 243  
937 TTGCTCTCTTTGAAGAGTGCAGCCACCCGCTGATGAAGATGAAGCAGCTGAGCAGC 996  
244 TTTGTCGGGAAGTATTCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGGCAAG 303  
997 TTTGTCGAGAAGTATTCGACACCATAGCTGAGTTGCGGAGCTGCGAGCCGTGGCGAGA 1056  
304 GACTTCGAAGTCAGAAGTCTTTGAGTTGTTGTTGCTGAGTGCAGCTGAGTGGTGAAGA 363  
1057 GACTTCGAAGTTTCGAAGCTTTGCGGCTGTGTTGCTCTGCTGAGTGCAGCTGAGTTAGA 1116  
364 GAGAAAGCAACCGGGGACATCTATGCTATGAAAGTGTAGGAAGAAAGAGGCTTTATGCCC 423  
1117 GAGAAGCGACCGGGGACGCTATGCTCATGAAATCATGAAGAAGAAGGCTTTGCTGGCC 1176  
424 CAGGAGAGGTTTCATTTTTCAGGAGAGCGGAGACATATTATCTCGAAGCACAAGCCCG 483  
1177 CAGGAACAGGTTTCATTTTTCGAGGAGGAGGAGAACATATTATCTCGGAGCAGAGTCT 1236  
484 TGGATCCCCCAATPACAGTATGCTCTTTCAGACACAAAATCACCTTTATCTGATGAGGAA 543  
1237 TGGATCCCCCAGTTACAGTACGCGCTTTTCAGGACAAAATAAACCTTTACCTGGTCATGAA 1296  
544 TATCAGCTCGAGGGGACTGCTGCTCACTTTTGAATAGATATGAGGACCACTAGATGAA 603  
1297 TATCAGCTCGAGGGGATTTCTGCTGCTCTGAAACAGATACGAGGACCAATTAGATGAG 1356  
604 AACCTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTTACACGGTTCATCTGATG 663  
1357 AGCATGATCAGTTTACCTTGTGCTGAGCTGATTTTGGCTGTCCACAGGTCACCATG 1416  
664 GGATACGTGATCAGACATCAAGCCCTGAGAACATTTCTGTTGACCCGACAGGACATC 723  
1417 GGATATGTGATCAGACATCAAGCCCGGAGAACATCTCATCGACCCGAGCGGAGAGATC 1476  
724 AAGCTGTGATTTTGGATCTGCCCGGAAATGAATTCACAAAGATGTTGAATGCCAAA 783  
1477 AAGCTGTGATTTTGGATCAGCCGCTAGATGAGTAAAGTAAATTCAAATAA ---GGTGGATGCCAAA 1533  
784 CTCCCGATTGGGACCCCGAGATTACATGGCTCTGAAAGTGTGATGTAACAGCGGAT 843  
1534 CTCCCGATTGGGACCCCGAGATTACATGGCTCTGAAAGTGTGATGTAACAGCGGAT 1593  
844 GGAAGGACACCTACGGCTGGAGTGTGATGAGTGTGCTGAGTGGCGGTGATTCCTATGAG 903  
1594 CGAAGGGGACACATACGGCTGGAGTGTGATGAGTGTGCTGAGTGGCGGTGATTCCTATGAG 1653  
904 ATGATTTATGGGAGATCCCCCTTCGAGAGGGAACCTTCGCGAAGACCTTCAATAAACAT 963  
1654 ATGTTTATGGGAAGACCCCATTCAGAGGGGAACCTTCGCGGAGACCTTCAACAACATC 1713  
964 ATGAATTTCCAGCGGTTTGAATAATTTCCAGATGACCCCAAGATGAGTACGATGCTTCTT 1023  
1714 ATGAATTTCCAGCGGTTTGAATAATTTCCAGATGACCCCAAGATGAGTACGATGCTTCTT 1773  
1024 GATCTGATTCAGAGCTTGTGTCGGCCAGAGAGAGAGATGGAAGTTTGAAGGCTTCTTGC 1083  
1774 GATCTGTTTCAAGTCTGTGCTGTGTCAGGAAGAAGAGAGATGGAAGTTTCGAGGGTCTCTGC 1833  
1084 TGCCATCCTTTCTTCTCTAAATTTGACTGGAAACAATTCGTAACCTCTCTCTCCCCCTTC 1143  
1834 TGCCACCCCTTTCTTTGCGAGAACCGACTGGAAACAACATCCGTAACCTCTCTCTCCCCCTTC 1893  
1144 GTTCCACCCCTCAAGTCTGAGGATGACACCTCCCAATTTGATGTAACAGAGAGAAATTCG 1203

ORIGIN

Query Match 70.5%; Score 4632.4; DB 10; Length 6954;  
Best Local Similarity 85.6%; Pred. No. 0;  
Matches 5341; Conservative 0; Mismatches 641; Indels 254; Gaps 7;  
4 GCCGCCAGTGGGAGATGTTGAAGTTCAAAATATGAGCGCGGGAATCCTTTGATGCTGCT 63  
757 GGGTTGGCGGAGAGATGTTGAAGTTCAAGTATGTTGTCGGGAACCGCGGAGGCGCAT 816  
64 GCTGCTGAACCCATTCGACCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCA 123

Db 1894 GTCCCCACCCCTCAAGTCTGACGATGACACCTCCAAATTTTGTATGAAACAGAGAAAGATTG 1953  
Qy 1204 TGGG-TTTTATCTCTCCGTCAGCTGAGCCCTCAGGTTCTCGGGTGAAGAACTGCC 1262  
Db 1954 TGGGCTTTTATCTCTCTGTGTGCCAGCTGAGCCCCCTC-GCGTTCTCAGGCGAAGAGCTGCC 2012  
Qy 1263 GTTTGTGGGGTTTTCGTACAGCAGGACCTGGGATTTCTTGGTAGTCTCTGAGTCTCTT 1322  
Db 2013 GTTTGTGGGATTTTCGTACAGCAGGACCTGGGATTTCTTGGTAGTCTCTGAGTCTCTG 2072  
Qy 1323 GTCCGGTCTGGACTCCCTCCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAG 1382  
Db 2073 GTCGAGTCTGGACTCCCTCCCAAGGTTAGCTCCATGGAAAAGAACTTCTCATCAAAG 2132  
Qy 1383 CAAGAGCTACAGACTCTCAGNCAAGTCTCAAGATGAGCAGGAATGACCCGGTT 1442  
Db 2133 CAAGAGCTCCAGACTCCCAAGGACAGTGTCAAGATGAGCAGGAATGACCCGGTT 2192  
Qy 1443 ACATCGGAGAGTGTCAAGGTGGAGGCTGTGCTTAGTCAGAAGGAGGTGAGCTGAAGGC 1502  
Db 2193 ACATCGCAGAGTGTCAAGGTGGAGGCTGTGCTTAGTCAGAAGGAGGTGAGCTGAAGGC 2252  
Qy 1503 CTCTGAGACTCAGAGATCCCTCTCGAGCAGGACCTTGTCTACTCATCAGAAATGCGAG 1562  
Db 2253 CTCTGAGACTCAGAGATCCCTCTCGAGCAGGACCTTGTCTACTCATCAGAAATGCGAG 2312  
Qy 1563 TAGCTTAAAGCGAAGTTTGAGCAAGCAGGATGGAGGTGCTCCAGAGGAGTGCACAAAGC 1622  
Db 2313 TAGCTTAAAGCGAAGTTTGAGCAAGCAGGATGGAGGTGCTCCAGAGGAGTGCACAAAGC 2372  
Qy 1623 ACTGAGCTTCTCATGATATCAGAGCAGAGCCCGAAGCTCCAGAAATCAAAGAGCA 1682  
Db 2373 TCTGAGCTTCTCCACGACATCCGAGAGCAGAGCCCGAAGCTCCAGAGATCAAAGAGCA 2432  
Qy 1683 GGAGTACAGGCTCAAGTGAAGAAATGAGTTGATGATGAATCACTGCGAGAGATCT 1742  
Db 2433 GGAGTACAGGCTCAGTGGAGGAGATGAGCTGATGAATCAGCTGAAGAGACCT 2492  
Qy 1743 TGTCTCAGCAAGAGAGCGAGTGTCTCTACGAATCTGAGCTGAGAGTCTCGGCTTGC 1802  
Db 2493 GGTGTACGCCGAGCAGCGACGATCTCTACGAGTCTGAGCTGAGGAGTCTCGGCTTGC 2552  
Qy 1803 TGTGAAGAAATCAAGCGGAAACGACAGAAATGTCTAGCATAACTTTGAAGCTTAAGGA 1862  
Db 2553 CGCCGAGGAATCAAGCGGAAACGAAATGTCTAGCAAACTGATGAAGCTTAAGGA 2612  
Qy 1863 TCAAGGGAAGCTTGAAGTGGGAGAAATGCGAACTGGAGAAGATCAATGCTGAGCAGCA 1922  
Db 2613 CCAAGGGAAGCTTGAAGTGGGAGAAATTTCCAACTGGAGAGATCAATGCTGAGCAGCA 2672  
Qy 1923 GCTCAAAATTCAGAGCTCCAGAGAAATCTGAGAAAGGCT- 1962  
Db 2673 GCTGAAGATCCAGAGCTCCAGGAGAGCTGGAGAAGGCTGTAAAGCCAGCAGCAGAGGC 2732  
Qy 1963 -----GCAAGGAGCGAGCGGAGAGGAGCTCGAGAA 1994  
Db 2733 CACCGAGCTCTCGAATATCCGCCAGGCAAGGAGCAGAGCGGAGCTTAGAA 2792  
Qy 1995 GCTGCAAGAACGAGAGGATTTCTTCAAGGCATCAAGAAAGAGCTGGTGAAGCTGAGGA 2054  
Db 2793 GCTTACACAAACCGGGAAGCTCTCCAGGGCATCAAAAGAAAGCTGTTGAAGCCGAGGA 2852  
Qy 2055 ACCCCGCCATCTCTGAGAAACAGGTAAAGAGCTAGAGCCATGAGCGCTAGAGAA 2114  
Db 2853 ACCCCGCCATCTCTGAGAAACAGGTAAAGAGCTAGAGCAATGAGAGCTGAGAGAA 2912  
Qy 2115 CAGACTGAAGGATGACATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATAAAAT 2174  
Db 2913 CAGACTGAAGGATGACATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATAAAAT 2972  
Qy 2175 TCTGGAGCTCGAGAGAAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGA 2234  
Db 2973 TCTGGAGCTCGAGAGAAACATCGGAGGCCCTCAGGCTCAACATCTAGAAAGTACA 3032

Qy 2235 CCTGAAAACAGAAAGAGCAGCACTATGAGGAAAAGATTAAAGTTGTGGCAATCAGATAAA 2294  
Db 3033 CTTGAAAACAGAAAGAAACAGCACTACGAGGAAAAGATCAAGGTATTGGCAATCAGATAAA 3092  
Qy 2295 GAAACACCTGCTGACAAAGGAGACACTCGAACAATGATGACAGACACGAGGAGGAGGC 2354  
Db 3093 GAAAGACCTGCTGACAAAGGAGAGCCTGGAGAAATGATGACAGACACGAGGAGGAGGC 3152  
Qy 2355 CCATCAGAAGGGCAAAATTTCTCAGCGAACAGAAAGCCGATGATCAATGCTATGGATTCCAA 2414  
Db 3153 CCACGAGAAGGGCAAGATCTCTCAGCGAGCAAGAGCGATGATCAAGCGATGGATTCCAA 3212  
Qy 2415 GATCAGATCCCTGGNAACAGAGGATTGTGGAATCTCTGGAAGCCAAATAAATTGACAGCAA 2474  
Db 3213 GATCGATCCCTGGAGCAGAGGATCGTGGAGCTGTGGAAGCCAAAGCTTGGCGGCAAA 3272  
Qy 2475 TAGCAGTCTTTTACCCAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAG 2534  
Db 3273 CAGCAGTCTCTTACCACAGAGGACATGAAGGCCCAAGAGAGATGATCTCAGAACTCAG 3332  
Qy 2535 GCAACAGAAATTTTACCTGGAGACACAGGCTGGGAAAGTTGGAGGCCCAAGAACCCGAAACT 2594  
Db 3333 GCAGCAGAAATTTTACCTGGAGACGAGCGCGGAAAGCTGGAGGCCCAAGAACCCGAAAGCT 3392  
Qy 2595 GGAGGAGCAGCTGAGAGAGATCAGCCAAAGACCAAGTGCACAGTGCACAAAGATCGGCTGTGGA 2654  
Db 3393 GGAAGAGCAACTGGAGAGATCAGCCCAAGATCAAGTGACAAAGAGTCCGCTGTGGA 3452  
Qy 2655 ACTGAGACAAAGATTGCGGAGGTGAGTCTAGAGCAAGAGAGCAGAAACTGAGAGCTCAA 2714  
Db 3453 GCTGAAAACAGGCTGAGGAGGTGAGCTGAGAGCAAGAGGAGCAGAACTGAGAGCTGAA 3512  
Qy 2715 CGCCAGCTCACAGCTACAGCTCTCCCTGACAGAGCGGAGTGCACAGTTGACAGCCCT 2774  
Db 3513 CGCGAGCTCAGGAGCTGAGCTGTCCCTGAGAGCGCGAGTCCCGCTGAGCTGACGGCCCT 3572  
Qy 2775 GCAGGCTGCAGCGGCGCCCTGGAGAGCCAGCTTCGCGAGCGGAAAGACAGAGCTGGAAGA 2834  
Db 3573 GCAGCGCCCGGCGAGCGCTGGAGAGCCAGCTCCGCGAGGAAAGACAGAGCTGGAGGA 3632  
Qy 2835 GACCACAGCAGAGCTGAAAGAGGATCAGGCACTCAGCGCACATAGATGAATCCA 2894  
Db 3633 GACAAACCGGAGAGGAGGAGATCCAGGCGCTCAGGCGCATCGCGATGAGATCCA 3692  
Qy 2895 GCGCAAAATTTGATGCTCTTCTGTAACAGCTGATGTAAATCACAGACCTGGAGAGCAGCT 2954  
Db 3693 GCGCAATTCGATGCCCTTCGCAACAGCTGCACGGTCACTACCGACTGGAGAGCAGCT 3752  
Qy 2955 AAACACAGCTGACCGAGAGCAACCGTGAATCTCAACAACTTTCTACTTTGTCGAAACA 3014  
Db 3753 GAAACAGCTCACCGAGAGCAACCGCGAGCTCAACAACTTTCTACTTTGTCGAAACA 3812  
Qy 3015 ACTCGATGAGCTTCTGCGCCCAACGACGAGATTGTACAACTGCAAGTGAAGTGAACCA 3074  
Db 3813 ACTCGATGAGCTTCTGCGGCCCAATGACGAGATTGTGACGCTGCGAGAGTGAAGNCCA 3872  
Qy 3075 TCTCCGCGGAGATCAGGAACGAGAGATGCGAGTTTACCGCCAGAAAGCAAAACGATGGA 3134  
Db 3873 TCTCCGCGGTGAGATCAGGAGCGGAGATGCGAGTCCACGAGCAGAAAGCAAAACGATGGA 3932  
Qy 3135 GGCTCTGAAGCAGCAGCTGACCATGCTGGAGGAGCAGGTCATGATTTGGAGCCCTTAA 3194  
Db 3933 GGCTCTGAAGCAGCAGCTGACCATGCTGGAGGAGCAGGTCCTGAACTGGAGCCCTTAA 3992  
Qy 3195 CGATGAGCTCTGAAGAAAGAGCGGAGTGGAGGCTGGAGGAGGCTCTGGGTGATGA 3254  
Db 3993 CGACGAGCTGCTGAGAAAGAGCGCAGTGGAGGCTGGCGAGCGTCTTGGCGACGA 4052  
Qy 3255 GAATCCAGCTTGTGAGTGTGCGGTTCGAGAGCTGAGAGATGCTGACACCGAGAAACA 3314  
Db 4053 GAATCCAGCTTGTGAGTGTGCGGTTCGAGAGCTGAGAGATGCTGACACCGAGAAACA 4112

QY 3315 GAGCAGGCGAGAGCGGATCAGCGGATCACCAGTCTCGCCAGGTGGTGGAGCTGGCAGT 3374  
DB 4113 GAGCAGGCTAGGGCCGATCAGCGGATCACCAGTCTCGCCAGGTGGTGGAGTGGCGGT 4172  
QY 3375 GAAGGACGACAGGCTGAGATTCTCGCTCTGAGCAGAGCTCTCAAGAGCAGAGCTGAA 3434  
DB 4173 GAAGGACGACAGGCTGAGATTCTCGCTCTGAGCAGAGCTCTCAAGAGCAGAGCTCAA 4232  
QY 3435 GCGCAGAGCCCTCTCTGACAAGCTCAATGACCTGGAGAGAGAGTGTATCTTTGAAAT 3494  
DB 4233 AGCCGAGAGCCCTCTCGGACAAGCTCAAGCTCAAGCTGGAGAGAGAGCAGCCATCTGGAGAT 4292  
QY 3495 GAATGCCGAGCTTTACGACGAGAGCTGGAGACTGAAACGAGAGCTCAACACAGAGCTTCT 3554  
DB 4293 GAACGCCGAGCTTTACGACGAGAGCTGGAGACTGAAACGAGAGCTCAACACAGAGCTTCT 4352  
QY 3555 GGAAGAGCAAGCCAAATTAACGACGAGATGACCTGCAAGAAATCAATTTTCGGTCT 3614  
DB 4353 GGAAGAGCAAGCCAAATTAACGACGAGATGACCTGCAAGAAATCAATTTTCGGTCT 4412  
QY 3615 GACTAAGGATGCAAGAGCTTTAGATCGGCTGATCTACTGAGAGCAAGAGAGTGA 3674  
DB 4413 GACGCAAGGCTGACGAGGCGCTGAGCGGCGGATCTGCTGAAGACAGAGAGAGGGA 4472  
QY 3675 CTGGAGTATCAGCTGGAAGCAATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGA 3734  
DB 4473 CTGGAGTATCAGCTGGAAGCAATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGA 4532  
QY 3735 AGGCACTATTTCTCAACAAACCAAACTCATTTGATTTCTGCAAGCCAAATGGACCAACC 3794  
DB 4533 AGGCACTATTTCTCAACAAACCAAACTCATTTGATTTCTGCAAGCCAAATGGACCAACC 4592  
QY 3795 TGTCTAAAAAGAAAAGGGTTATTTAGTTCGAGCGGAAAGAGGACCTGCTTTACCCACACA 3854  
DB 4593 TGTCTAAAAAGAAAAGGGTTATTTAGTTCGAGCGGAAAGAGGACCTGCTTTACCCACACA 4607  
QY 3855 GGTTCCTCTGCACTACAATGAGCTGAAGCTGGCCCTGGAGAGAGGAGGAGGCTCGCTGTGC 3914  
DB 4608 GGTTCCTCTGCACTACAATGAGCTGAAGCTGAGCCCTGGAGAGAGGAGGAGGCTCGCTGTGC 4667  
QY 3915 AGAGCTAGAGGAGGCTCTGAGAGGACCCGATTCGAGCTCGGCTCGCCCGGAGGAGG 3974  
DB 4668 GGAGCTGGAGGAGGCTCTGAGAGGACCCGATTCGAGCTCGGCTCGCCCGGAGGAGG 4727  
QY 3975 TGCCCAACCGCAAGCAGGACCAACCCACCCATCCAGCCAGCCAGCCAGCCAGCCAGCA 4034  
DB 4728 TGCCCAACCGCAAGCAGGACCAACCCAGCCAGCCAGCCAGCCAGCCAGCCAGCA 4787  
QY 4035 GATCGCATGTCTGCGCATCTGCGGTTCGCGAGGACCCAGCCAGGAGGAGGAGGAGG 4094  
DB 4788 GATCGCATGTCTGCGCATCTGCGGTTCGCGAGGACCCAGCCAGGAGGAGGAGGAGG 4847  
QY 4095 GGCCCGCATCCAGCGCGAGAGAGGAGTCTCAACTCAGAGGAGTTCAGTTCGGGCTCT 4154  
DB 4848 TGCCCAACCGCAAGCAGGAGGAGTCTCAACTCAGAGGAGTTCAGTTCGGGCTCT 4907  
QY 4155 TAAGGAACGATGCACCAATATTCCTCACCAGTTCACAGTTCAGGAGTTCAGGAGTTC 4214  
DB 4908 GAAGAGCGGATGCACCAATATTCCTCACCAGTTCACAGTTCAGGAGTTCAGGAGTTC 4967  
QY 4215 CACAAAGT 4274  
DB 4968 CACAAAGT 5027  
QY 4275 CGAATGTTCAGT 4334  
DB 5028 CGAATGTTCAGT 5087  
QY 4335 GCCTGTGTAATGCCACACACTTCACCGAGGCTTCTGCGGTGACAAATGAACTCCCC 4394  
DB 5088 GCCAGCTGTAATGCCACACACTTCACCGAGGCTTCTGCGGTGACAAATGAACTCCCC 5147  
QY 4395 AGGTCTCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4454

DB 5148 GGGCTCCAGCAGGAGGAGCTGGCAGCAGCTTGCACCTGGAAGGGTGGATGAAAGTGCC 5207  
QY 4455 CAGGAATAAACAACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4514  
DB 5208 CAGGAATAAACAACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 5267  
QY 4515 AAAAGTCTCTATTTATGACAAATGAGCCAGAGAGCTGGACAGAGGCGGTGGAGAGATT 4574  
DB 5268 AAAAGTCTCTATTTATGACAAATGAGCCAGAGAGCTGGACAGAGGCGGTGGAGAGATT 5327  
QY 4575 TGAGCTGTGCTTCCCGCAGCGGGATGATCTATTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4634  
DB 5328 TGAGCTGTGCTTCCCGCAGCGGGAGCTATCTATTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5387  
QY 4635 CGCAAAATACAGCCAAAGCA----- 4653  
DB 5388 TGCAAAATACCGCCAAAGCAGATGTCCCATATCCTGAAGATGGAGTCTCACCCACACAC 5447  
QY 4654 ----- 4653  
DB 5448 CACCTGTGGCTGGGAGAACCTCTACTTGTGTAGCACCAGCTTCCCGCAGCAGCAGCG 5507  
QY 4654 -----GAAAAAGCAGA 4664  
DB 5508 CTGGGTCAACCGCTTAGAATCTGCTGCGCAGGTGGAGAGATTCTTAGGAAAGAGCCGA 5567  
QY 4665 AGCTGATGCTTAACTGCTTGGAAACTCCTGCTGTAACCTGGAAGCTGATGACCTCTAGA 4724  
DB 5568 AGCCGATGCTAAATTAATCTTGGAACTCTCTGCTGTAACCTGGAAGCTGATGACCTCTAGA 5627  
QY 4725 CATGAATGTCAGCTGCGCTTCACTGACCAAGTGGTGTGTTGGTGGCACCAGAGAGGCT 4784  
DB 5628 CATGAATGTCAGCTGCGCTTCACTGACCAAGTGGTGTGTTGGTGGCACCAGAGAGGCT 5687  
QY 4785 CTACGCTGTAATGCTTGTGAAAACTCCCTAAACCTGCTCCAGGAATTTGGAGCAGTCTT 4844  
DB 5688 CTACGCTGTAATGCTTGTGAAAACTCCCTAAACCTGCTCCAGGAATTTGGCAGTCTT 5747  
QY 4845 CCAATTTATATTTCAAGGACCTTGAAGAGTCTCATGATAGCAGAGAGAGGCGGC 4904  
DB 5748 CCAATTTATATTTCAAGGACCTTGAAGAGTCTCATGATAGCAGAGAGAGGCGGC 5807  
QY 4905 ACTGTGTCTTGGACGTGAAGAAAGTGAACAGTCTGCTGCGCCAGTCCCACTTCTGCTGC 4964  
DB 5808 TCTGTGCTGTGGAGCTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 5867  
QY 4965 CGAGCCGACATCTCAACCAATTTTGAAGCTGTCAGAGGCTGCAAGGCTGCAAGGCTGCAAGGCT 5024  
DB 5868 CGAGCCGACATCTCAACCAATTTTGAAGCTGTCAGAGGCTGCAAGGCTGCAAGGCTGCAAGGCT 5927  
QY 5025 AGCAAGATTGAGAAAGGCTCTGCACTGTGACAGCCTATGCGCCAGTCCCACTTCTGCTGCT 5084  
DB 5928 TGGCAAGATCAGAGACAGCTTGTCACTGTGCGCTATGCAAGCAAGTCTGCTGCTGCT 5987  
QY 5085 CGCTACAAAGAAACCTCAGCAAAATCTGATCGGAAAGAGATAGAGACTCAGAGCT 5144  
DB 5988 CGCTACAAAGAAACCTCAGCAAAATCTGATCGGAAAGAGATAGAGACTCAGAGCT 6047  
QY 5145 CTGAGCTGTATCCACTTCAACCAATTTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAG 5204  
DB 6048 CTGAGCTGTATCCACTTCAACCAATTTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAG 6107  
QY 5205 AATCGACATGAAGCAGTACAGCTCGAGGAATTCCTGGATGAAGATGACCAATTCCTTGGC 5264  
DB 6108 GATCGACATGAAGCAGTACAGCTCGAGGAATTCCTGGATGAAGATGACCAATTCCTTGGC 6167  
QY 5265 ACCTGCTGTGTTCGGCTTCTTCAACAGCTTCCCTGCTCAATTCGTCAGAGGAGGAGCAG 5324  
DB 6168 ACCTGCTGTGTTCGGCTTCTTCAACAGCTTCCCTGCTCAATTCGTCAGAGGAGGAGCAG 6227  
QY 5325 CGCAGGCGAGGAGGAGTACTTGTGTGTTCACCAATTTGGAGTTCGTCGTCGTCGTCGTCGTCG 5384



Db 910 ATGATGAATCAGCTGGAAGAGGACCTGCTGTGTCAGCCCGAAGACGAGCGATCTCTACGAG 969  
QY 1777 TCTGAGCTCAGAGAGTCTCGGCTTGCTGCTGAAGAATTCAACGGAAACGACAGAAATGT 1836  
Db 970 TCTGAGCTGAGGAGTCTCGGCTCGCGCCGAGGAATTCAGCGGAAGCAAAACGAATGT 1029  
QY 1837 CAGCATAAACTTTGAAAGCTTAAGGATCAAGGAGAGCTGTAAGTGGGAGATATCGGAA 1896  
Db 1030 CAGCATAAACTGATGAAGCTTAAGGACCTAGGGAAGCTGAAAGTGGGAAATGTTCGAGA 1089  
QY 1897 CTGGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAGCTCAAGAGAAATCTGGAG 1956  
Db 1090 CTGGAGAAGATCAATGCTGAGCAGCAGCTGAAGATTCAGGAGCTCCAGGAGAAGCTGGAG 1149  
QY 1957 AAGGCT-----GCAAG 1968  
Db 1150 AAGGCTGTAAAGCCAGCACCGAGGCCACAGAGCTCTCTGAGAAATCCGGCAGGCAAG 1209  
QY 1969 GAGCAGCCGAGGAGGAGCTGAGAGCTGCGAGAACTCGAGAGAGGATTTCTTGAAGGCATC 2028  
Db 1210 GAGCAGCTGAGGAGGAGCTGAGAGCTGCAACACAGGGAAGATTTCTCGAAGGCATC 1269  
QY 2029 AGAAGAAGCTGGTGAAGCTGAGGAACCGCGCATTTCTTGAGAAACAAAGTAAAGAGA 2088  
Db 1270 AAAAGAAAATCTGTGAAGCTGAGGAACCGCGCACTCCCTGGAGAAACAAAGTAAAGAGA 1329  
QY 2089 CTAGAGACCATGAGCGTAGAGAAACACACTGAAGGTGACATCCAGACAAATCCCAA 2148  
Db 1330 CTAGAGACCATGAGCGTAGAGAAACACACTGAAGGTGACATCCAGAAAGTCCGAA 1389  
QY 2149 CAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTGGAAGAGAAACATCGGAGGCCAA 2208  
Db 1390 CAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTGGAGGAGAAACATCGGAGGCTCAG 1449  
QY 2209 GTCTAGCCCGACCATGAGTGAAGTGCACCTGAAACAGAGAGCAGACTATCAGAGAAAG 2268  
Db 1450 GTCTAGCTCAACACTGAAGTACACTGAAGCAGAGAGAAACAGACTATCAGAGAAAG 1509  
QY 2269 ATTAAAGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGACAAGGAGACACTGGAGAAC 2328  
Db 1510 ATCAAGTGTGGACAAATCAGATAAAGAGGACTTGGCCGACAGAGAGCCTAGAGACC 1569  
QY 2329 ATGATCCAGAGACAGAGAGAGGCGCATGAGAGGGCAAAATTTCTAGCGAAACAGAG 2388  
Db 1570 ATGATCCAGAGGCCAGAGAGAGGCGCCACGAGAGGGCAAGATCCTCAGCGAGCAGAG 1629  
QY 2389 GCGATGATCAATGCTATGATTCAGATCCAGATCCCTGGAACAGAGGATTTGTGAACTG 2448  
Db 1630 GCGATGATCAACGCAATGGAATTCAGATCCCTGGAAGAGGATTTGTGAGCTG 1689  
QY 2449 TCTGAAGCCAAATAAATTCAGCAAAATAGCAGTCTTTTACCCAAAGGAAACATGAAGGCC 2508  
Db 1690 TCAGAAGCCAAACAAGCTTGGGCAACAGCAGTCTCTCCACCCAGAGGAAACATGAAGGCC 1749  
QY 2509 CAAGAGAGATGATTTCTGAATCAGGCAACAGAAATTTTACCTGGAGACAGCTGGG 2568  
Db 1750 CAGGAAGAGATGATCTCAGAACTCAGGCGAGCAAAATTTTACCTGGAGACAGCGCCGG 1809  
QY 2569 AAGTTGGAGGCCAGAACCGAAATCTGGAGGAGCAGCTCGAGAAGATCAGCCACCAAGAC 2628  
Db 1810 AAGCTGGAGGCCAGAACCGAAATCTGGAAGACACTAGAGAGATCAGCCACCAAGAT 1869  
QY 2629 CACAGTGACAGAAATCGGCTGTGAACTCGGAGACAAAGATTGGGAGAGTCACTAGAG 2688  
Db 1870 CACAGTGACAAAGATCGGCTGTGAGCTGAAGCGGACGCTCAGCGAGCTGCAGCTGTCCCTG 1929  
QY 2689 CAGCAGGAGCAGAAATCGGAGCTCAAGCGCCAGCTCAGAGAGTACAGCTCTCCCTGCGAG 2748  
Db 1930 CATGAGAGCAGAACTGAGCTGAAGCGGACGCTCAGCGAGCTGCAGCTGTCCCTGCGAG 1989  
QY 2749 GAGCGGAGTCAAGTGAAGCCCTGCGAGGCTGCAAGCGCGCCCTCGAGAGCGAGCTTT 2808  
Db 1990 GAGCGTGAAGTCCAGCTGAGCGCCCTGCGAGCGCGCCCTGCCCTTGAGAGCGAGCTC 2049

QY 2809 CGCCAGCGCAGAGACAGAGCTGGAAAGAGACCAACAGCAGAGAGCTGAAGAGGAGATCCAGGCA 2868  
Db 2050 CGCCAGCGCAGAGACAGAGCTGGAGGAGCAACATCTCAGAGCGGAGGAGGATCCAGGCA 2109  
QY 2869 CTACAGCACAATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTCTAAACAGCTGTACT 2928  
Db 2110 CTACAGCACAATCGAGATGAATCCAGCGCAAAATTCGATGCCCTTCGCAACAGCTGTACT 2169  
QY 2929 GTAATCAGACCTGGAGGAGCAGCTAAACCAAGCTGACCGAGGACAACGCTGAATCAAC 2988  
Db 2170 GTCATCACCGACCTGGAGGAGCAGCTGAACCAAGCTCAACGAGGACAACGCGAGCTCAAC 2229  
QY 2989 AACCAAACTTTCTACTTTGTCACAACTCGATGAGGCTTCTGGCGCAACAGCAGAT 3048  
Db 2230 AACCAAACTTTCTACTTTGTCACAACTCGATGAGGCTTCGGGTGCAACGATGAT 2289  
QY 3049 GTACAACTGCGAAGTGAAGTGGACCATCTCCGCGGAGATCAACGGAACAGAGATCGAG 3108  
Db 2290 GTGAGCTGCGAAGCGAGGTGGACCATCTCCGCGTGAATCACGAGCGGAGATCGAG 2349  
QY 3109 CTTACCGCCAGAGCAAAACGATGGAGCTCTGAAGACCACTGTCACCATGCTGGAGAA 3168  
Db 2350 CTCACCGCCAGAGCAAAACGATGGAGCTCTGAAGACGACCTGCACCATGCTGGAGAG 2409  
QY 3169 CAGGTCATGGAATTTGGAGGCGCTAAACGATGAGCTGTAGAAAAGAGCGGCTGGGAG 3228  
Db 2410 CAGGTCATGAGCTTGGAGGCGCTGAAACGAGGCTGTGAGAGAGGAGCGCCAGTGGAG 2469  
QY 3229 GCCTGGAGGAGCTCTCGGTGATGAGAAATCCAGTTTGAAGTGTCCGGTTCGAGAGCTG 3288  
Db 2470 GCCTGGAGGAGCTCTCGGGGACGAAAGTCTCAGTTTGAAGTGTGAGTTCGAGAGCTA 2529  
QY 3289 CAGAGAACTGTGGACACCGAGAAACAGAGAGGCGGAGAGCCGATCAGCGATCAGCGAG 3348  
Db 2530 CAGAGAACTGTGGACACCGAGAAACAGAGAGGCTGAGAGCTGATCAGCGATCAGCGAG 2589  
QY 3349 TCTCGCCAGGTGGTGGAGCTGGCAGTGAAGAGACCAAGGCTGAGATTTCTGCTCTCGAG 3408  
Db 2590 TCTCGCCAGGTGGTGGAGCTGGCAGTGAAGAGACCAAAAGCCGAGATTTCTCGGCTCGAG 2649  
QY 3409 CAGGCTCTCAAGAGCAGAGAGCTGAAGGCGGAGAGCTCTCTGACAGCTCAATGACCTG 3468  
Db 2650 CAGGCTCTCAAGAGCAGAGAGCTGAAGAGCTGAGAGCTGTGGGACAAAGCTCAAGAGCTG 2709  
QY 3469 GAGAGAGCAGTGTATGCTTGAATGAATGCCCGAAGCTTACAGCAGAAAGCTGGAGACT 3528  
Db 2710 GAGAGAGCAGCAGCCTGATGATGAATGCCCGGAGCTTACAGCAGAACTGAGAGACA 2769  
QY 3529 GAAACGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAATTTACAGCAGCAGATGGAC 3588  
Db 2770 GAGCGGAGCTCAACAGAGGCTTCTGGAGGAGCAAGCCAAAGTTGCGAGCAGCAGATGGAC 2829  
QY 3589 CTGCAGAAAATACATTTTCCGCTCTGACTCAAGGACTGCAAGAGCTCTAGATCGGGCT 3648  
Db 2830 CTGCAGAAAGAACCATCTTCAGACTGACACAGGCTGCGAGGCGCTGGACCGGCG 2889  
QY 3649 GATCTAAGAGACAGAAAGTGAATTTGAGATCTGAGTATCAGCTGGAAGAAACATTCAGGTTCTC 3708  
Db 2890 GATCTGTGAAGACCGAAAGAGGCGACTGGAATATCAGCTGGAAGAAACATTCAGGTTCTC 2949  
QY 3709 TATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAACAAACCAAACTCATTTGAT 3768  
Db 2950 TACTCACAGAGAGGTGAAATGGAAGGCACTATTTCTCAACAAACCAAACTCATTTGAT 3009  
QY 3769 TTTCTGCAAGCCAAATGGACCAACCTGCTAAAAAGAAAAAGGGTTTATTTAGTCGACGG 3828  
Db 3010 TTTCTGCAAGCCAAATGGACCCGCTAAAAAGAAAAAGGGTTTATTTAGTCGACGG 3069  
QY 3829 AAAGAGAGCCTGCTTTACCCACAGAGTTCTCTGAGTACATGAGCTGAAGCTGCC 3888  
Db 3070 AAAGAGAGCCTGCTTTGCCACACAGGTTCTCTGCAATGAGCTGAAGCTGCC 3129

QY 3889 CTGGAGAGGAGAAAGCTCGTGTGCAGAGCTAGAGGAAGCCCTTCAGAGAGACCCGCATC 3948  
DB 3130 CTGGAAAGAGGAAAGCGCGCTGTGCGAGCTGGAGGAAGCCCTTCAGAGAGACCCGCATC 3189  
QY 3949 GAGCTCCGGTCCGCCCGGAGGAGCTGCCACCGCAAGCAACCGGACCAACCCACACCCA 4008  
DB 3190 GAACTCCGGTCCGCACCGGAGGAGCTGCTCACCGCAAGCAACCGGACCAACCCACACCCA 3249  
QY 4009 TCCACGCCAGCCACCGGAGGAGCAGATCGCCATGTCGCCCATCTGCGGTCCGAG 4068  
DB 3250 TCTACGCCAGCCATCGGAGGAGCAGATCGCCATGTCGCCCATCTGCGGTCCGAG 3309  
QY 4069 CACAGCCCATGTCATGAGCCTGTGCGCCCGCCCATCCAGCCGCAAGAGAGTCTTCA 4128  
DB 3310 CACAGCCCATGTCATGAGCCTGTGCGCCCGCCCATCCAGCCGCAAGAGAGGCGTCA 3369  
QY 4129 ACTCAGAGGAACTTAGTCGGCTCTTAAGGAAGCAATGCACCAATATTCCTCACCGA 4188  
DB 3370 ACTCAGAGGATTCAGCCGGGCTTTAAAGAACGTATGCACCAACATTCCTCACCGG 3429  
QY 4189 TTCAACGTAGCACTGAACATGCCAGCCACAAAGTGTCTGTGTGGATACCGTGCAC 4248  
DB 3430 TTTAACTGGGCTGAATATGCGAGCCACCAAGTGGCGGTGTCTGGATATCTGTGCAC 3489  
QY 4249 TTGGAGCCAGGCAATCCAAATGTCTGGAATGTGAGGTGATGTGCACCCCAAGTCTCC 4308  
DB 3490 TTTGGAGCCAGGCAATCCAAATGTCTGGAATGTGAGGTGATGTGCCATCCCAAGTCTCC 3549  
QY 4309 ACGTGCTTGGCCAGCCACTCGCGCTTCCCTGCTGAATATGCCACACACTTCACCGAGGCC 4368  
DB 3550 ACTGCTTGGCCAGCCACTCGCGCTTCCCTGCTGAATATGCCACACACTTCACCGAGGCC 3609  
QY 4369 TTTGCGCGTGACAAATGAACTCCCGAGTCTCCAGCAAGGAGCCAGCAGCTTG 4428  
DB 3610 TTTGCGCGAGCAAAAGTGAAGTCTCCCGGCTCCAGCAAGGAGCCAGCAGCTTG 3669  
QY 4429 CACTGGAAGGTGGATGAGGTGCCAGGAAATACAAAGAGGAGCAGCAGCTGGAC 4488  
DB 3670 CACTGGAAGGTGGATGAGGTGCCAGGAAATACAAAGAGGAGCAGCAGCTGGAC 3729  
QY 4489 AGGAAGTACATTTCTCTGGAGGATCAAAAGTCTCTCATTTATGACAAATGAAGCCAGAA 4548  
DB 3730 AGAAGTACATTTCTTAGAGGTTCAAAAGTCTCTCATTTATGACAAACGAGCCAGAA 3789  
QY 4549 GCTGGACAGAGCGGTGGAGAAATTTAGCTGTGCTTCCCGACGGGATGATCTATT 4608  
DB 3790 GCTGGACAGAGCGGTGGAGAAATTTAGCTGTGCTTCCCGACGGGATGATCTATT 3849  
QY 4609 CATGGTCCGCTTGGTCTCCGAATCGCAATACAGCCAAAGCA----- 4653  
DB 3850 CATGGTCCGCTTGGTCTTCTGAATTCGAAATACAGCCAAAGCAGATGTCCTCATATA 3909  
QY 4654 ----- 4653  
DB 3910 CTGAAGATGAGTCCATCCGCATACCACTGCTGGCTGGAGAACCTCTACCTGCTG 3969  
QY 4654 ----- 4653  
DB 3970 GCACCCAGCTTCTTGACAAAGAGCGCTGGGTACAGCCCTAGAGTCTGTGTCGAGGT 4029  
QY 4654 -----GAAAGCAGAGCTGATGCT----- 4674  
DB 4030 GGGAGAGTTCTAGGGAAAGGAGGAGCGGATGCTGCAATGAGCACTGATCTTCTGTGAG 4089  
QY 4675 -----AACTGCTTGGAACTCCCTGCTGAACTGAACTGAAAGTGTAT 4713  
DB 4090 CGTCTGCCGTGGGTGTGAAATTTACTTGGAACTCTCTGTGAAGTGAAGTGTGAC 4149  
QY 4714 GACCGCTTACATGACATGACGCTGCCCTTTCAGTGACAGGTGGTGTGGTGGGCACC 4773  
DB 4150 GACCGCTTACATGACATGACGCTGCCCTTTCAGTGACAGGTGGTGTGGTGGGCACC 4209  
QY 4774 GAGGAAGGCTCTACGCCCTGAATGTCTTTGAAAACTCCCTAACCCATGTCCAGGAATT 4833

DB 4210 GAGGAGGCGCTCTATGCGCTGAACGCTTTGAAAACTCCTTAAACCATCCAGGAATC 4269  
QY 4834 GGAGCAGTCTTCCAAATTTATATTAAGGACCTGGAGAAAGCTACTCATGATAGCAGGA 4893  
DB 4270 GGAGCAGTCTTCCAGATTTACATCAAGGACCTGGAGAGCTGCTCATGATAGCAGGG 4329  
QY 4894 GAAGAGCGGCGACTGTGCTTGTGAGTGAAGAAAGTGAACACAGTCCCTGGCCAGTCC 4953  
DB 4330 GAAGAGCGGCGCTGTGTTGTGAGAGTGAAGAAAGTGAAGCAGTCTGTGGCAGCTCA 4389  
QY 4954 CACTGCTGCCAGCCGACATCTACCCCAACTTTTGAAGTGTCAAGGGTGGCCAC 5013  
DB 4390 CACTTCTGCCAGCGCTGAGCTCTCCCAACATAATTGAAGCGCTCAAAGGTGCCAC 4449  
QY 5014 TTGTTGGGCGAGCAGATGGAAGCGGCTCTGCACTCTGCGAGCATGCGCCAGCAAA 5073  
DB 4450 CTGTTGTGCTGGCAGATCGAAGCAGCCTCTGCACTCTGTGCCCATGCCAGCAAA 4509  
QY 5074 GTCGTCACTTCCGCTACACGAAACCTCAGCAAAATCTGCAATCCGAAAGAGATAGAG 5133  
DB 4510 GTCGTCACTTCCGCTACACGCAACCTCAGCAAGTCTGCACTCCGCAAGGATGAG 4569  
QY 5134 ACCTCAGAGCCCTGAGCTGTATCCACTTACCAATTTACAGTATCCTCATTGGNACCAAT 5193  
DB 4570 ACCTCAGAGCCCTGAGCTGTATCCACTTACCAATTTACAGTATCCTCATCGGAACCAAC 4629  
QY 5194 AAATCTACGAAATCGACATGAAGCAGTACACGCTCGAGGAATTCCTGGATAAGATGAC 5253  
DB 4630 AAGTTCTATGAGATCGACATGAACAGTACACGCTTGGAGTTCCTGGACAGATGAC 4689  
QY 5254 CATTCCTTGGCACCTGTGTGTTGCGCTCTTCCAAAGCTTTCCTGTCTCAATCGTG 5313  
DB 4690 CATTCCTTGGCACCTGTGTGTTGCGCATCTCTCGACCAACAGCTTCCCGCTCTCATCGTG 4749  
QY 5314 CAGGTGAACAGCGCAGCGCAGGAGAGTACTTGTGTTTCCACGAAATTTGGAGTG 5373  
DB 4750 CAGGAAACAGCACCGGCGAGCGGAAAGTACTCTGCTGTGTTTCCACGAAATTTGGG 4809  
QY 5374 TCTGTGATTTTACGGAAGACGTAGCCGACAGACGATCTCAAGTGGAGTTCGCTTACCT 5433  
DB 4810 TCTGTGATTTTATGAAGACGTAGCCGACAGACGATCTCAAGTGGAGTTCGCTTACCT 4869  
QY 5434 TTGCGCTTTCCTACAGAAACCTATCTGTTGTGACCCACTTCACTTCACTCGAGTA 5493  
DB 4870 CTGCGCTTCGCTACAGAGAGCTTATCTGTTGTGACTCACTTTAACTCCCTGGAAAGTA 4929  
QY 5494 ATTGAGATCCAGGACACGCTCTCAGCAGGAGACCCCTGCCGAGCGTACCTGGACATCCCG 5553  
DB 4930 ATTGAGATCCAGGACACGCTCTCAGTGGGACCCCTGCCGAGCTATCTGGAATCCCA 4989  
QY 5554 AACCCGCGCTACCTGGGCGCTGCCATTTCTCAGGAGGATTTACTTGGCGTCTCATAC 5613  
DB 4990 AACCTCGCTACCTGGGCGCGCAATTTCTCCGAGCAATTTACCTGGCGCTCTCATAC 5049  
QY 5614 CAGGATAAATAAGCGCTATTGCTGCAAGGAAACCTCTGGAAGGAGTCCGGCACTGAA 5673  
DB 5050 CAGGACAAAGTTAAGGTTCATATGCTGCAAGGAAACCTCTGGAAGGAGTCTGGCAGT 5109  
QY 5674 CACCAAGCGGCGCTCCAGCTCCCGCAGCAGCCCAACAAAGGAGGCGCCACCCAGCTAC 5733  
DB 5110 CAGCAGCGGTGCGCTCCACATCCC---GCAGCGCTTAAACAGGAGGCGCCACCAATAC 5166  
QY 5734 AACGAGCAGATCAACCAAGCGCTGGCTCCAGCCGAGCGCGCCGAGGCGCCAGCCAC 5793  
DB 5167 AACGAGCAGATCAACCAAGCGTGGCTCCAGCCGAGCGCCACGAGAGGCGCCAGCCAT 5226  
QY 5794 CCGGAGAGCAAGCAACACCCACCGCTACC-----GCAGGGGCGGACCGAGTGC 5847  
DB 5227 CCGGAGAGCAAGCAACACCCACCGCTACCGGAGAGAGGCGCGGACAGAGTGC 5286  
QY 5848 AGGAGCAAGTCTCTGCGCGCGCTGGAGCGAGAGTCCCGCGCGGATGCTCAGC 5907

Db 5287 AGGACAAAGTCTCAGGCGCTCTCTGGAGCGGAGAACTGCCAGCGGATGCTCAGC 5346  
Qy 5908 ACCGAGAGAGCGGTCCCGGAGGCTGTTTGAACAGACAGACAGCGGCGCGTGCCT 5967  
Db 5347 ACAGAGAGAGCGGTCCCGGAGGACTGTTTGAAGACAGACAGCGGCGCGTGCCT 5406  
Qy 5968 GCGGAGCGGTGAGGAGCGGCTGTCCTCCAGGTGAACAAGGAGAGGCGAGAGTGCCTCT 6027  
Db 5407 GCGGAGCGGTGAGGAGCGGCTGTCCTCCAGGTGAACAAGGCTTGGAGACAGTCTCAGTA 5466  
Qy 6028 CAGTTT 6034  
Db 5467 TAAGTCT 5473

RESULT 15  
AF070066  
LOCUS Mus musculus Citron-K kinase mRNA, partial cds. ROD 31-JUL-1998  
DEFINITION AF070066  
ACCESSION AF070066  
VERSION AF070066.1 GI:3360513  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 4967)  
Madaule, P., Fuyuyashiki, T., Reid, T., Ishizaki, T., Watanabe, G.,  
Morii, N. and Narumiya, S.  
A novel partner for the GTP-bound forms of rho and rac  
FEBS Lett. 377 (2), 243-248 (1995)  
96128238  
PUBMED 8543060  
REFERENCE 2 (bases 1 to 4967)  
Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuo, T., Bito, H.,  
Ishizaki, T. and Narumiya, S.  
Role of Citron kinase as a target of the small GTPase Rho in  
cytokinesis  
Nature (1998) In press  
REFERENCE 3 (bases 1 to 4967)  
Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuo, T., Bito, H.,  
Ishizaki, T. and Narumiya, S.  
Direct Submission  
Submitted (04-JUN-1998) Pharmacology, Faculty of Medicine, Kyoto  
University, Yoshida-Konohe, Sakyo-ku, Kyoto 606 8315, Japan  
FEATURES  
Location/Qualifiers  
1..4967  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="brain"  
/dev\_stage="adult"  
<1..4928  
/function="binds to Rho and regulates cytokinesis"  
/note="kinase variant form of Citron; approximately the  
first 50 amino acids belong to the kinase domain;  
alternatively spliced form of Citron, encoded by GenBank  
Accession Number AF039218, similar to Rattus norvegicus  
Citron-K encoded by GenBank Accession Number AF070065"  
/codon\_start=3  
/product="Citron-K kinase"  
/protein\_id="AAC27933.1"  
/db\_xref="GI:3360514"  
/translation="PEVPTLKSDDDTSNFDSEPKNSWSSVCSQSPSGFSGELPFV  
GRSYKALGYLGNSESVSSLDSPAKVSSMEKKLLIKKELQSDQKCHKMEQVTRL  
HRVSEAVLSQKVELKASFTQSLLEQDIATYITCCSLSKRSLEQARMEVSEDD  
KALQLLRIDRSKRLQETKEQYQAVEMRLMNLQNEEDLVARSRRSDYSELR  
SRLAEFFRKNECQHKLMKADQGPVGYSLKLEKINAEQQLKIQLQKLEKAV  
KASTELLQNRQAKERAERELEKHNRESSEGIKKLVEASELEKREAOVSA  
OHLVELKQEOHYEKKIQLDNOIKKDLADKESLENMQRHEEERHKKGLISQKA  
MINAMSKTSLSQRIVELSEANKLAANSLETFQNMKAQEMISELRQKQYLETQA  
KLEAFNRKLEGLKIRSHQSDRSKRLLETRVSELEHEERKLEKRLQTLLEQL  
SLQERESQJLQAAPAALESQLRQAKTELETTABEIEIQAHTAHRDEIQKQKFDAL

QY 1856 CTAAGGATCAAGGAGAGCTGAAGTGGAGATATGCGAACTGGAGAGATCAATGCTG 1915  
Db |||||  
QY 721 CTAAGGACCAAGGAGAGCTGAAGTGGAGATATCCAACTGGAGAGATCAATGCTG 780  
Db |||||  
QY 1916 AGCAGAGCTCAAAATTCAGGAGCTCAAGAGAACTGGAGAGGCTGCAAGAGCGAG 1975  
Db |||||  
QY 781 AGCAGAGCTCAAGATCCAGAGCTCAGAGAGCTGGAGAGGCTGTAAAGCGAGCA 840  
Db |||||  
QY 1976 CCGAGAGGAGCTGGAGAGCTCGAGAACCGAGAGGATTTCTGAGGCGATCAGAAAGA 2035  
Db |||||  
QY 841 CAGAGGCCACCGAGCTCCTGAGAAATATCCCGCAGG-----CAAAGG 882  
QY 2036 AGCTGGTGGAGCTGAGAGAACCGCGCATCTCTGGAGAACAGGTAAGAGACTAGAGA 2095  
Db |||||  
QY 883 AGCAGAGAGCGGAG-----CTAGAGAGCTTACAAACCGGGAAGACT 927  
QY 2096 CCATGGAGCTAGAGAAACAGACTGAGGATGACATCCAGACAAATCCAAACAGATCC 2155  
Db |||||  
QY 928 CCTCCGAGGGCATCAAAAGAGCTGTGTGAAGC----- 962  
QY 2156 AGCAGATGGCTGATAAAATTCGGAGCTCGAAGAGAACATCGGAGGCCCAAGTCTCAG 2215  
Db |||||  
QY 963 -----GAGAGCTGGAGAGAGAACATCGGAGGCTCAGGTCTCAG 1002  
QY 2216 CCAGACCTAGAGTGCACCTGAAACAGAGAGAGAGCAGCTATGAGGAGAAAGATTAAAG 2275  
Db |||||  
QY 1003 CTCAACATCTAGAGTACACTTGAACAGAGAGAACAGCACTACGAGGAGAAAGATCAAGG 1062  
QY 2276 TGTGGACATCAGATTAAGAGAGCTGCTGCTGACAGAGAGACACTGGAGAACATGATGC 2335  
Db |||||  
QY 1063 TATGGACATCAGATTAAGAGAGCTGCTGCTGACAGAGAGAGCTGGAGAACATGATGC 1122  
QY 2336 AGAGACAGAGGAGGAGGCTGAGAGAGGCAAAATTTCTCAGCGAACAGAGGCCATGA 2395  
Db |||||  
QY 1123 AGAGACAGAGGAGGAGGCTGAGAGAGAGGCAAGATCTCAGCGAGAGAGGCCATGA 1182  
QY 2396 TCAATGCTATGGATTCAGAGATCAGATCCCTGAGAACAGAGGATTTGGAGACTGCTGAAG 2455  
Db |||||  
QY 1183 TCAACGGATGGATTCAGATCCGATCCCTGGAGCAGAGGATCTGGAGCTGTCGGAAG 1242  
QY 2456 CCAATRAACTGAGCAAAATAGGAGCTTTTACCAAGAGAACATGAGGCCCAAGAG 2515  
Db |||||  
QY 1243 CCAACAGCTTGCGGCAACAGCAGTCTTCCACAGAGAGAACATGAGGCCCCAGAGAG 1302  
QY 2516 AGATGATTTCTGAACCTCAGGCAACAGAAATTTTACCTGGAGACACAGGCTGGAGAGTTGG 2575  
Db |||||  
QY 1303 AGATGATCTCAGAACTCAGGACAGAGAAATTTTACCTGGAGAGCAGGCGGGAAGCTGG 1362  
QY 2576 AGCCCGAGACCCGAAACTGGAGGAGCAGCTGGAGAGATCAGCCACAGACCAAGTGG 2635  
Db |||||  
QY 1363 AGGCCAGAACCCGAAAGCTGGAAGAGCACTGGAGAGATCAGCCACCAAGATCAGAGT 1422  
QY 2636 ACAAGATCGGCTGCTGGAACCTGGAGACAGATGCGGAGGCTCAGTCTAGAGCAGAGG 2695  
Db |||||  
QY 1423 ACAAGAGCTCGCTGAGCTGGAACAGGCTGAGGAGGCTCAGCTGGAGCAGAGG 1482  
QY 2696 AGCAGAACTGGAGCTCAAGCGCAGCTCAGAGAGTACAGTCTCCTGCGAGAGCGCG 2755  
Db |||||  
QY 1483 AGCAGAGCTGGAGCTGGAACCGCAGCTCAGGAGCTGAGCTGCTCCTGCGAGAGCGCG 1542  
QY 2756 AGTCACAGTTGACAGCCCTCAGGCTGACCGGCGGCTGGAGAGCCAGCTTCGCGAGG 2815  
Db |||||  
QY 1543 AGTCCAGCTGACGGCCCTGACGCCCGCGGACGCTGGAGAGCAGCTCCGCCAGG 1602  
QY 2816 CGAAGACAGAGCTGGAGAGAGACACAGAGAGAGCTGAGAGAGATCCAGGCACTCAGG 2875  
Db |||||  
QY 1603 CGAAGACAGAGCTGGAGAGAGAACCGGAGAGAGAGAGATCCAGGCGCTCAGCG 1662  
QY 2876 CACATGAGATGAATCCACGCAAAATTTGATGCTCTCTGTAACAGCTGCTACTGTAATCA 2935  
Db |||||  
QY 1663 CACATCGGATGAGATCCAGCGCAATTCGATGCCCTTGGCAACAGCTGCACGGTATCA 1722  
QY 2936 CAGACCTGGAGAGAGCTTAAACAGCTGACGAGAGCAACCGTGAATCTAAACAAACAAA 2995

Db |||||  
QY 1723 CCGACTGGAGAGAGCTGAACCAAGCTCACCGAGAGCAACCGCAGCTCAACAAACAAA 1782  
QY 2996 ACTTCTACTTTTCCAAACAACTCGATGAGGCTTCTGGGCCCAACGACGAGATTGTACAAC 3055  
Db |||||  
QY 1783 ACTTCTACTTGTCAAAACAACTCGATGAGGCTTCGGGGCCCAATGACGAGATTGTGCAAG 1842  
QY 3056 TCGAAGTGAAGTGAACCATCTCCGCCGGAGATCACGAAGCAGAGATGAGCTTACCA 3115  
Db |||||  
QY 1843 TCGGAGTGAAGTGAACCATCTCCGCCGTGAGATCACGAGCGGAGATGAGCTACCA 1902  
QY 3116 GCGAAGCAAAACGATGAGGCTCTGAAGACCACTGCAACCATGCTGGAGAAACAGGTCA 3175  
Db |||||  
QY 1903 GCCAGAACAAACGATGAGGCTCTGAAGACGACATGACCATGTTGGAGAGCAGGTCC 1962  
QY 3176 TGGATTTGGAGGCCCTTAAACGATGAGCTGTAGAAAAGAGCGGAGTGGAGGCTGGA 3235  
Db |||||  
QY 1963 TGGACTGGAGGCCCTGAAACGACGAGCTGTGGAGAGGAGCGCAGTGGAGGCTTGGC 2022  
QY 3236 GGAGGCTCTCGGTGATGAGAAATCCCAAGTTTGTAGTGTGGGTTTCGAGAGCTGCAGAGAA 3295  
Db |||||  
QY 2023 GGAGGCTCTTGGCGACGAGAAATCCCAAGTTTGTAGTGTGGGTTTCGAGAGCTGCAGAGAA 2082  
QY 3296 TGTGGACACCGAGAAACAGAGCAGGCGAGCGGATCAGCGGATCACCGAGTCTGGCC 3355  
Db |||||  
QY 2083 TGTGGACACCGAGAAACAGAGCAGGCGTGGGCGGATCAGCGGATCACCGAGTCTGGCC 2142  
QY 3356 AGGTGGTGGGCTGGAGTGAAGGACACAGGCTGAGATTTCTGCTGTGAGAGGCTC 3415  
Db |||||  
QY 2143 AGGTGGTGGGCTGGAGTGAAGGAAACAAAGGCGGAGATTTCTGCTGTGAGAGGCTC 2202  
QY 3416 TCAAGAGCAGAGCTGAAGCCGAGAGGCTCTCTGACAACTCAATGACCTGGAGAGAA 3475  
Db |||||  
QY 2203 TCAAGAGCAGAGCTCAAGCCGAGAGGCTCTCGGACAACTCAACGCTCAGAGAGAA 2262  
QY 3476 AGCATGCTATCTTGAATGAATGCCGAGGCTTACAGCAGAGCTGGAGACTGAAAGAG 3535  
Db |||||  
QY 2263 AGCAGCGCATCTGGAGATGAACCGCGGAGCTTACAGCAGAACTAGAGACAGAGCGGG 2322  
QY 3536 AGCTCAACAGAGGCTCTTGAAGAGCAAGCAAAATTTACAGCAGCAGATGGACCTGAGAG 3595  
Db |||||  
QY 2323 AGCTCAACAGAGGCTCTTGGAGAGCAGAGCAAAATTTGAGCAGCAGATGGACCTGAGAG 2382  
QY 3596 AAAATCACTTTTCCGCTGCTGACTCAAGGATGCAAGAGCTCTAGATCGGCTGATCTAC 3655  
Db |||||  
QY 2383 AGAACCACATCTTCAGACTGACGCAAGGCTGAGAGGCGCTGGACCGGCGGATCTGC 2442  
QY 3656 TGAAGACAGAGAGAGTGAAGTGAAGTATCAGCTGGAGAAACATTCAGGTTCTCTATTCTC 3715  
Db |||||  
QY 2443 TGAAGACAGAGAGAGGAGCCTGGAGTACAGCTGGAGAAACATTCAGGTTCTCTACTCTC 2502  
QY 3716 ATGAAAAGGTGAAAATGGAAGGCACTATTTCTCAACAAACCAAACTCATTCATTGTTTCTGC 3775  
Db |||||  
QY 2503 ACAGAGAAAGTGAAGGCACTCTCTCAGCAAAACCAAACTCATTCATTGTTTCTGCTGC 2562  
QY 3776 AAGCCAAATGGAACCACTCTTAAGAGAAAGGCTTTATTAGTCGAGGAAAGAGG 3835  
Db |||||  
QY 2563 AAGCCAAATGGAACCACTCTTAAGAGAAAGGCTTTATTAGTCGAGGAAAGAGG 2595  
QY 3836 ACCCTGCTTTACCCACAGAGTCTCTGCACTCAATGAGCTGAAGCTGGCCCTGGAGA 3895  
Db |||||  
QY 2596 -----AGGTTCTCTGCACTCAATGAGCTGAAGCTAGCCCTGGAAG 2637  
QY 3896 AGGAGAAAGCTCGCTGTGAGAGCTAGAGAGAGCCCTTCAGAGAGCCGCGATCGAGGCTCC 3955  
Db |||||  
QY 2638 AGGAGAAAGCCGATCGCGAGCTGGAGAGGCGCTTCAGAGAGCCCGCATCGAATCC 2697  
QY 3956 GGTCCGCGCGGAGAGAGCTGCCCAACGCAAGCAACGAGCAGCCACCCACCTCCACGC 4015  
Db |||||  
QY 2698 GCTCTGCGCGGAGAGAGTGGCCCAACGCAAGCAACGAGCAGCCACCCGCTCTACGC 2757  
QY 4016 CAGCCACCGGAGCAGAGATGCGCATGTCCGCCATCTGCGGTCGCGAGAGCAGCAGC 4075

Db 2758 CAGCCACTGCGAGGACAGATCGCCATGTCTGCCATTTGTGGTTCGCCGCGAGCACCAGC 2817  
Qy 4076 CCAGTGCCTAGCCTGCTGGCCCGCCATCCAGCCGCGAGAAAGAGTCTTCAACTCCAG 4135  
Db 2818 CAGTGCCTAGCCTGCTGGCCCGCCATCCAGCCGCGAGAAAGAGTCTTCAACTCCAG 2877  
Qy 4136 AGGAATTTAGTGGCGCTTTAAGGAACGATGACACACCAATATTCCTCACCGAATTCACG 4195  
Db 2878 AGGAATTCAGCGCGCTGTGAAGAGCGATGCACCACAACATCCCTCACCGGTTAATG 2937  
Qy 4196 TAGGACTGAACATCGAGCCACAAAGTGTCTGTGTCTGTATACCGTGCACCTTGGAC 4255  
Db 2938 TGGGCTGAACATGAGAGCCACAAGTGGCCGCTGTGTGTGTATCTGTGCACCTTGGAC 2997  
Qy 4256 GCCAGCATCCAAATGCTCGAATGTCAGGTGATGTGTCAACCAAGTGTCTCCAGTGTCT 4315  
Db 2998 GCCAGCATCCAAATGCTCGAATGTCAGGTGATGTGTCAACCAAAATGCTCCACCTGCT 3057  
Qy 4316 TGCAGCCACTTGGCGCTTGCCTGTGTAATATGCCACACACTTCCAGGAGGCTTCTGCC 4375  
Db 3058 TGGCCGCTACTGTGGCCTGCCAGCTGAATATGCCACACACTTCACTGAGGCGCTTCTGCC 3117  
Qy 4376 GTGACAAAATGAATCCCAAGTCTCCAGACCAAGGAGCCAGCAGCTTGCACCTGG 4435  
Db 3118 GGGACAAAATGAATCTCCCGGGCTCCAGAGCAAGAGCTGCGACGACCTTGCACCTGG 3177  
Qy 4436 AAGGTGGATGAAGTGGCCAGGAATPACAAACGAGACAGCAAGCTGGGACAGGAAGT 4495  
Db 3178 AAGGTGGATGAAGTGGCCAGGAATPACAAACGAGGACAGCAAGCTGGGACAGGAAGT 3237  
Qy 4496 ACATGTCTGGAGGATCAAAAGTCTCATTTATGACAAATGAACATGAAGCAGAGTGGAC 4555  
Db 3238 ACATTTCTGGAGGGGTCAAAAGTCTCATTTATGACAAATGAAGCAGAGAGTGGAC 3297  
Qy 4556 AGAGGCGGTGGAAGAAATTTAGCTGTGCTTCCGAGCGGGATGTATCTATTCTATGGTG 4615  
Db 3298 AGAGGCGGTGGAAGAAATTTAGCTGTGCTTCCGAGCGGACGATCTATCTATCTATGGTG 3357  
Qy 4616 CGTTGGTGGTTCGAACTCGCAATATACGCAAAAGCA----- 4653  
Db 3358 CGTTGGTGGTTCGAACTTCGAAATACCGCCAAAGCAGATGTCCCATATCCTCGAAG 3417  
Qy 4654 ----- 4653  
Db 3418 TGGAGTCTCACCCACACACACCTGCTGGCTGGGAGAACCTCTACTTGTAGCACCCA 3477  
Qy 4654 ----- 4653  
Db 3478 GCTTCCCGCACAAGCAGCGCTGGGTCAACCGCTTAGAACTCTGTGTGGAGGTGGAGAG 3537  
Qy 4654 -----GAAAGCAGAGCTGATGCTAAACTGCTTGGAACTCCCTGCTGAACCTGG 4705  
Db 3538 TTTCTAGGAAAGGCGGAGCGGATGCTAAATTACTTGGAACTCTCTGCTGAACCTGG 3597  
Qy 4706 AAGGTGATGACCGCTTAGACATGAATGACGCTGCTTCAAGTACCAAGTGGTGTGG 4765  
Db 3598 AGGCGATGACCGGCTTGACATGAATGACACCTGCTTCAAGTACCAAGTGGTGTGG 3657  
Qy 4766 TGGGCAACGAGGAGGCTTAGCGCTGAATGCTTGAAGAACTCCCTTAACCCATGTGCC 4825  
Db 3658 TGGGCAACGAGGAGGCTTAGCGCTGAATGCTTGAAGAACTCCCTTAACCCATGTGCC 3717  
Qy 4826 CAGGAATGGAGCAGTCTTCCAAATTTATATCAAGGACCTGGAGAGCTACTCATGA 4885  
Db 3718 CAGGATTTGGCGAGTCTTCCAAATTTATCATCAAGGACCTGGAGAGCTGTCTCATGA 3777  
Qy 4886 TAGCAGAGAGAGCGGCACTGTGCTTTGTGGAAGTGAAGAAAGTGAACAGTCCCTGG 4945  
Db 3778 TAGCAGGGAAGAGCGGCTCTGTGCTGTGGACGTGAAGAAAGTGAAGCAGTCCCTGG 3837  
Qy 4946 CCAGTCCCACTGCTGCTCCAGCCGACATCTCACCCCAACATTTTGAAGTGTCAAG 5005  
Db 3838 CGCAGTCAACCTGCTGCTCCAGCCGAGCTCTCCCCCAACATATTCGAAGCGGTCAAG 3897

Qy 5006 GCTGCCACTTGTGGGCGAGGCAAGATTGAGAACGGGCTCTGCATCTGTGACGCCATGC 5065  
Db 3898 GCTGCCACTTGTGGGCGAGGCAAGATTGAGAACGGGCTCTGCATCTGTGACGCCATGC 3957  
Qy 5066 CCAGCAAGTGTCTCATTTCTCGCTTACAACGAAACCTCAGCAATACTGATCCGGAAG 5125  
Db 3958 CAAGCAAGTGTCTCATCTCTCGCTTACAATGACAACTCAGCAAGTACTGATCCGCAAG 4017  
Qy 5126 AGATGAGACCTCAGAGCCCTGAGCTGTATCCACTTCAACAAATCAGATCTCTCATTTG 5185  
Db 4018 AGATGAGACCTCAGAGCCCTGAGCTGTATCCACTTCAACAAATCAGATCTCTCATTTG 4077  
Qy 5186 GAACCAATAAATTTACGAAATCGACATGAAGCAGTACAGCTCGAGGAATTCCTGGATA 5245  
Db 4078 GGACCAAAATTTATGAGATGACATGAAGCAGTACAGCTGTGATGAGTTCCTGGACA 4137  
Qy 5246 AGAATGACATTTCTTGGCACCCTGCTGTGTGGCGCTCTTCCAAAGCTTCTCTCTCT 5305  
Db 4138 AGAAGCACTTCTTGGCACCCTGCTGTGTGGCGCTCTTCCAAAGCTTCTCTCTCTCT 4197  
Qy 5306 CAATCGTGCAGGTGAACAGCGCAGGCGAGGAGAGTACTTGTCTGTGTTCACGAAAT 5365  
Db 4198 CCATGTGTCAGGCGAAACAGCGCCGCGCAGGAGAGTACTTGTCTGTGTTCACGAAAT 4257  
Qy 5366 TTGAGTGTTCGTGGATTTCTTAACGAAAGCTAGCCGACAGAGATCTCAAGTGGAGTC 5425  
Db 4258 TTGGGTGTTCGTGGATTTCTTAACGAAAGCTAGCCGACAGATGATCTCAAGTGGAGTC 4317  
Qy 5426 GCTTACCTTGGCCTTTGCTTACAGAAACCTTATCTGTGTGTGACCCACTTCAACTCAC 5485  
Db 4318 GCTTACCTTGGCCTTTGCTTACAGAAACCTTATCTGTGTGTGACTCATCTCAACTCCC 4377  
Qy 5486 TCGAAGTAATAGATCCAGCACGCTCTCAGCAGGAGCCCTGCGCAGCGTACCTGG 5545  
Db 4378 TGAAGTCAATAGATCCAGGCCAGATCTCCTCTGGGAGCCCTGCGCAGCATATCTGG 4437  
Qy 5546 ACATCCGAAACCCGCTTACCTGGGCGCTGCTTCTCAGGAGCCATTTACTTGGCGT 5605  
Db 4438 AAATTCGAAACCCGCTTACCTGGGCGCTGCTTCTCAGGAGCCATTTACTTGGCGT 4497  
Qy 5606 CCTCATACCGAGTAATAAGGTCTTGTCTCAAGGGAACCTCTGCTGAAGGAGTCCG 5665  
Db 4498 CCTCATACCGAGTAATAAGGTCTTGTCTCAAGGGAACCTCTGCTGAAGGAGTCCG 4557  
Qy 5666 GCATGTAACACCGGCGCGCTTCTCCTCCGACAGCCGCGCAACAGGAGGCGCAC 5725  
Db 4558 GCATGTAACACCGGCGCGCTTCTCCTCCGACAGCCGCGCAACAGGAGGCGCAC 4617  
Qy 5726 CCAGTACAACGAGCACATCAACGCGGTGGCTCCAGCCAGCGCCCGCCGAGGCC 5785  
Db 4618 CAATATACACGAGCACATCAACGCGGTGGCTCCAGCCGCGCGCACCGGAGGCC 4677  
Qy 5786 CCAGCCACCCGAGAGCCAGCACACCCCAACCGCTACC-----GCGAGGGCGGACCG 5839  
Db 4678 CCAGCCACCCGAGAGCCAGCACACCGCACCGCTACCAGACAGAGAGGGTCCGACAG 4737  
Qy 5840 AGCTGCCAGGAGCAAGTCTCTGGCGCGCCCTTGGAGCGAGAGTCTCCCGGCGCGA 5899  
Db 4738 AGCTGCCAGGAGCAAGTCTCTGGCGCGCCCTTGGAGCGGAGTCTCCCGGCGCGA 4797  
Qy 5900 TGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTTGTGGAAGACAGCAGCAGGCGCC 5959  
Db 4798 TGCTCAGCAGCTAGGAGAGCGGTCCCGCGGAGTGTGGAAGACAGCAGCAGGCGCC 4857  
Qy 5960 GGCTGCTGGGAGCGGTGAGGACCCCGCTGCTCCAGGTGAACAGAGGAGAGGCGAGA 6019  
Db 4858 GGCTGCTGGGAGCGGTGAGGACCCCGCTGCTCCAGGTGAACAGAGTCTGGGACCACT 4917  
Qy 6020 GTGCTCTCAAGTTT 6034  
Db 4918 CTTAGTATAGTCT 4932

Mon Jul 26 15:37:17 2004

us-10-017-216-1.rge

Page 48

Search completed: July 4, 2004, 03:42:04  
Job time : 24322 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2004, 09:03:14 ; Search time 396 Seconds  
(without alignments)  
9212.745 Million cell updates/sec

Title: US-10-017-216-1  
Perfect score: 6574  
Sequence: 1 agagcgccagtgaggagat.....atcgagaatgagtttaga 6574

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/PTCUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1397.8	21.3	1515	4	US-09-804-471A-1
2	1397.8	21.3	1515	4	US-10-238-709-1
3	768	11.7	1133	4	US-09-916-204-1
4	258	3.9	258	4	US-09-016-434-513
5	220.8	3.4	2726	2	US-08-422-699A-12
6	220.8	3.4	2726	2	US-08-422-705B-12
7	214	3.3	2511	2	US-08-422-699A-8
8	214	3.3	2511	2	US-08-422-705B-8
9	213.6	3.2	3182	1	US-08-484-044-11
10	205	3.1	174493	4	US-09-804-471A-3
11	205	3.1	174493	4	US-10-238-709-3
12	189.2	2.9	2706	2	US-08-630-822A-61
13	189.2	2.9	2706	2	US-09-005-069-61
14	189.2	2.9	2706	4	US-09-171-156A-20
15	189.2	2.9	2706	4	US-09-004-730A-20
16	189.2	2.9	2706	4	US-08-981-799A-20
17	181.2	2.8	4363	2	US-08-685-576-5
18	179.6	2.7	4848	4	US-09-976-594-295
19	178.8	2.7	48763	4	US-09-916-204-3
20	172.2	2.6	4065	4	US-09-016-434-1105
21	172.2	2.6	4739	3	US-08-685-871-1
22	168.4	2.6	5053	2	US-08-685-576-2
23	150.2	2.3	3323	2	US-08-422-699A-10
24	150.2	2.3	3323	2	US-08-422-706B-10
25	122.6	1.9	3155	4	US-09-442-100-7
26	122.6	1.9	3155	4	US-08-939-106-7
27	122.6	1.9	3155	4	US-09-442-102-7

28	122.2	1.9	1498	4	US-09-509-902A-6	Sequence 6, Appli
29	122.2	1.9	1961	4	US-09-509-902A-15	Sequence 15, Appli
30	122.2	1.9	5276	4	US-09-233-857-2	Sequence 2, Appli
31	117.8	1.8	3018	2	US-08-850-150-6	Sequence 6, Appli
32	117.8	1.8	3018	3	US-09-338-132-6	Sequence 6, Appli
33	111.4	1.7	1935	2	US-08-878-989-11	Sequence 11, Appli
34	111.4	1.7	1935	3	US-09-272-796-11	Sequence 11, Appli
35	109.8	1.7	3213	4	US-09-442-100-5	Sequence 5, Appli
36	109.8	1.7	3213	4	US-08-939-106-5	Sequence 5, Appli
37	109.8	1.7	3213	4	US-09-442-102-5	Sequence 5, Appli
38	105	1.6	3489	2	US-08-728-223A-1	Sequence 1, Appli
39	105	1.6	3489	4	US-09-298-568-1	Sequence 1, Appli
40	105	1.6	3489	4	US-09-410-399-1	Sequence 1, Appli
C 41	105	1.6	32207	2	US-08-770-379-20	Sequence 20, Appli
C 42	105	1.6	32207	3	US-08-757-669A-20	Sequence 20, Appli
C 43	105	1.6	32207	4	US-09-230-371A-20	Sequence 20, Appli
44	104.8	1.6	2160	3	US-09-588-256-1	Sequence 1, Appli
45	101.4	1.5	5720	4	US-09-442-100-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-804-471A-1  
; Sequence 1, Application US/09804471A  
; Patent No. 6479269  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: C0601164  
; CURRENT APPLICATION NUMBER: US/09/804,471A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1515  
; TYPE: DNA  
; ORGANISM: Human  
US-09-804-471A-1

Query Match		21.3%	Score 1397.8;	DB 4;	Length 1515;
Best Local Similarity		99.5%	Pred. No. 0;		
Matches 1402;		Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Qy	13	GGGAGATGTTGAAGTTCAAATATGGAGCGCGGAATCTTTGGATGCTGTGCTGCTGAA	72		
Db	1	GGGAGATGTTGAAGTTCAAATATGGAGCGCGGAATCTTTGGATGCTGTGCTGCTGAA	60		
Qy	73	CCCATTCGCTGAGTCCGGGCTCCAGCTGAATCTCTTCCAGGGGAAACACCCCTTATG	132		
Db	61	CCCATTCGCTGAGTCCGGGCTCCAGCTGAATCTCTTCCAGGGGAAACACCCCTTATG	120		
Qy	133	ACTCAACAGAGATGCTCTCTTTCCGAGAGAGGATATTAGATGCCCTCTTTGTTCTC	192		
Db	121	ACTCAACAGAGATGCTCTCTTTCCGAGAGAGGATATTAGATGCCCTCTTTGTTCTC	180		
Qy	193	TTTGAAGATGTCAGTCAAGCTGCTCTGTGAAGATTAAGCACTGAGCACTTTGCCG	252		
Db	181	TTTGAAGATGTCAGTCAAGCTGCTCTGTGAAGATTAAGCACTGAGCACTTTGCCG	240		
Qy	253	AAGTATTCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGCGAAAGGACTTCGAA	312		
Db	241	AAGTATTCGACACCATAGCTGAGTTACAGAGCTTCGCGAAAGGACTTCGAA	300		
Qy	313	GTGAGAAGTCTTGTAGTTGTGCTACTTTGTGAAGTGCAGTGTGTAAAGAGAAGCA	372		
Db	301	GTGAGAAGTCTTGTAGTTGTGCTACTTTGTGAAGTGCAGTGTGTAAAGAGAAGCA	360		
Qy	373	ACCGGGACATCTTCTATGAAGTGAAGCAAGCAAGCTTTATGCCCCAGGACG	432		



QY 793 GGGACCCAGATTACATGGCTCTCTGAAGTGTGACTGTGATGAACGGGGATGGAAAAGGC 852  
DB 781 GGGACCCAGATTACATGGCTCTCTGAAGTGTGACTGTGATGAACGGGGATGGAAAAGGC 840  
QY 853 ACCTAGCGCTGGAGTGTGACTGTGCTGAGTGGGCTGATTCCTATGAGATGATTTAT 912  
DB 841 ACCTAGCGCTGGAGTGTGACTGTGCTGAGTGGGCTGATTCCTATGAGATGATTTAT 900  
QY 913 GGGAGATCCCCCTTCGCGAGAGGAAACCTCTGCGAGAACCTTTCAATAACATTATGAATTC 972  
DB 901 GGGAGATCCCCCTTCGCGAGAGGAAACCTCTGCGAGAACCTTTCAATAACATTATGAATTC 960  
QY 973 CAGCGGTTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1032  
DB 961 CAGCGGTTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020  
QY 1033 CAAAGCTTGTGTCGGCCAGAAAGAGAGACTGAAAGTCTTTGAGTGGCATCCT 1092  
DB 1021 CAAAGCTTGTGTCGGCCAGAAAGAGAGACTGAAAGTCTTTGAGTGGCATCCT 1080  
QY 1093 TTCCTCTCTAAATTTGACTGGAAACATTCGTAACTCTCTCCCTTCCTGTTCCAC 1152  
DB 1081 TTCCTCTCTAAATTTGACTGGAAACATTCGTAACTCTCTCCCTTCCTGTTCCAC 1140  
QY 1153 CTCAGTCTGAGATGACACCTCCCAATTTGATGAACAGAGAAATTCGTTGGTTCA 1212  
DB 1141 CTCAGTCTGAGATGACACCTCCCAATTTGATGAACAGAGAAATTCGTTGGTTCA 1200  
QY 1213 TCCTCTCGCTGACGCTGAGCCCTCAGCCTTCTCGGTTGAAGTCTCGGTTGGGG 1272  
DB 1201 TCCTCTCGCTGACGCTGAGCCCTCAGCCTTCTCGGTTGAAGTCTCGGTTGGGG 1260  
QY 1273 TTTTGTACAGAGCACTGGGATCTTGTGTAGATCTGAGTCTGTTGTGCGGCTG 1332  
DB 1261 TTTTGTACAGAGCACTGGGATCTTGTGTAGATCTGAGTCTGTTGTGCGGCTG 1320  
QY 1333 GACTCCCTGCAAGCTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTA 1392  
DB 1321 GACTCCCTGCAAGCTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTA 1380  
QY 1393 CAAGACTCTCAGCAAGTGTCAAGAT 1421  
DB 1381 CAAGACTCTCAGCAAGTGTCAAGAT 1409

RESULT 3  
US-09-916-204-1  
; Sequence 1, Application US/09916204  
; Patent No. 6638745  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001164CIP  
; CURRENT APPLICATION NUMBER: US/09/916,204  
; CURRENT FILING DATE: 2001-07-24  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1133  
; TYPE: DNA  
; ORGANISM: Human  
US-09-916-204-1

Query Match 11.7%; Score 768; DB 4; Length 1133;  
Best Local Similarity 99.4%; Pred. No. 7.3e-197;  
Matches 771; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGAGCGCCGAGTGGGAGATGTTGAAGTTCAATATGAGCGCGGAATCCCTTTGATGCT 60  
DB 36 AGAGCGCCGAGTGGGAGATGTTGAAGTTCAATATGAGCGCGGAATCCCTTTGATGCT 95

QY 61 GGTGCTCTGAACCCCAATTGCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAA 120  
DB 96 GGTGCTCTGAACCCCAATTGCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAA 155  
QY 121 CCACCCCTTTATGACTCAACAGCAGATGTCCTCTTCCGAGAGGGATATTAGATGCC 180  
DB 156 CCACCCCTTTATGACTCAACAGCAGATGTCCTCTTCCGAGAGGGATATTAGATGCC 215  
QY 181 CTCTTTGTTCTTTGAAGAAATCAGCTCAGCTCTCTCTGATGAAGATTTAGCAGCTGAGC 240  
DB 216 CTCTTTGTTCTTTGAAGAAATCAGCTCAGCTCTCTCTGATGAAGATTTAGCAGCTGAGC 275  
QY 241 AACTTTGTCGGGAAGTATTCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCA 300  
DB 276 AACTTTGTCGGGAAGTATTCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCA 335  
QY 301 AAGGACTTCCGAAGTCAGAACTCTTTGAGTTGTTGCTCAGCTTTGCTGAAGTGCAGGTGGA 360  
DB 336 AAGGACTTCCGAAGTCAGAACTCTTTGAGTTGTTGCTCAGCTTTGCTGAAGTGCAGGTGGA 395  
QY 361 AGAGAGAAACCAACCGGGGACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTG 420  
DB 396 AGAGAGAAACCAACCGGGGACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTG 455  
QY 421 GCCCAGGAGCAGGTTTCATTTTTTGGAGAGAGCGGAACATATTTATCTCGAAGCACAAGC 480  
DB 456 GCCCAGGAGCAGGTTTCATTTTTTGGAGAGAGCGGAACATATTTATCTCGAAGCACAAGC 515  
QY 481 CCGTGGATCCCCCAATTACAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAG 540  
DB 516 CCGTGGATCCCCCAATTACAGTATGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAG 575  
QY 541 GAATATCAGCTCGAGGGGACCTTCTGCTACATTTTGAATAGATATGAGGACCAAGTTAGAT 600  
DB 576 GAATATCAGCTCGAGGGGACCTTCTGCTACATTTTGAATAGATATGAGGACCAAGTTAGAT 635  
QY 601 GAAACCTGATACAGTTTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCTG 660  
DB 636 GAAACCTGATACAGTTTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCTG 695  
QY 661 ATGGGATACGTGATCGAGACATCAAGCTCGAGAAATCTCTGTTGACCGCCAGAGCAC 720  
DB 696 ATGGGATACGTGATCGAGACATCAAGCTCGAGAAATCTCTGTTGACCGCCAGAGCAC 755  
QY 721 ATCAAGCTGGTGGATTTTGGATCTGCCGGAATAATGAATTTCAACAAGATGGTGA 776  
DB 756 ATCAAGCTGGTGGATTTTGGATCTGCCGGAATAATGAATTTCAACAAGATGGTGA 811

RESULT 4  
US-09-016-434-513  
; Sequence 513, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Sellhammer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 513:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 258 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAINON01  
CLONE: 2290031  
US-09-016-434-513

Query Match 3.9%; Score 258; DB 4; Length 258;  
Best Local Similarity 100.0%; Pred. No. 1.2e-59;  
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5069 GCAAGTCGTATTCCTCCGCTACACGAAACCTCAGCAAACTCGCATCCGGAAGAGA 5128  
DB 1 GCAAGTCGTATTCCTCCGCTACACGAAACCTCAGCAAACTCGCATCCGGAAGAGA 60  
QY 5129 TAGAGACCTCAGAGCCCTCGAGCTGTATCCACTTACCAATACAGTATCCTCATTTGAA 5188  
DB 61 TAGAGACCTCAGAGCCCTCGAGCTGTATCCACTTACCAATACAGTATCCTCATTTGAA 120  
QY 5189 CCAATAAATCTACGAAATCGACATGAAGCAGTACACGCTCAGGAATTCCTGGATAGA 5248  
DB 121 CCAATAAATCTACGAAATCGACATGAAGCAGTACACGCTCAGGAATTCCTGGATAGA 180  
QY 5249 ATGACATTCTTGGACCTGCTGTGTTGCGCCCTCTTCCACAGCTTCCCTGTCTCAA 5308  
DB 181 ATGACATTCTTGGACCTGCTGTGTTGCGCCCTCTTCCACAGCTTCCCTGTCTCAA 240  
QY 5309 TCGTCAGGTGAACAGCG 5326  
DB 241 TCGTCAGGTGAACAGCG 258

RESULT 5  
US-08-422-699A-12  
Sequence 12, Application US/08422699A  
Patent No. 5985265  
GENERAL INFORMATION:  
APPLICANT: Brook, J. David  
APPLICANT: Housman, David E.  
APPLICANT: Shaw, Duncan J.  
APPLICANT: Harley, Helen G.  
APPLICANT: Johnson, Keith J.  
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC  
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,699A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/422,706  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/023,612  
FILING DATE: 26-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/839,255  
FILING DATE: 20-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/01545  
FILING DATE: 19-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/00253  
FILING DATE: 05-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB9202485.0  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: MIT-5830A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2726 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-422-699A-12  
Query Match 3.4%; Score 220.8; DB 2; Length 2726;  
Best Local Similarity 53.7%; Pred. No. 6e-49;  
Matches 531; Conservative 0; Mismatches 442; Indels 15; Gaps 3;  
QY 212 CTGCTCTGATGACATTAAAGCAGCTGAGCAACTTGTCCGGAAGTATTCGACACCATAG 271  
DB 243 CCGAACTGGCCCGAGGCAAGTACGTGGCCGACTTCTTGAGTGGCGGAGCCATCGTG 302  
QY 272 CTGAGTTACAGGAGCTCCAGCCTTCGCAAAAGGACTTCGAAAGTCAGAAAGTCTTTGTAGGTT 331  
DB 303 TGAGGCTTAAGGAGGTCGACTCGAGAGGACGACTTCGAGATTCTGAAGGTGATCGGAC 362  
QY 332 GTGGTCACTTTGCTGAAGTCCAGTGGTAGAGAGAGCAACCGGGGACATCTATGCTA 391  
DB 363 GCGGGCGCTTCAGCGAGGTAGCGGTAGTGAAGATGAAGCAGACGCGGCCAGGTGTATGCCA 422  
QY 392 TGAAGTGTATGAAGAAAGGCTTTATTGGCCCGAGGAGGCTTTTCATTTTTTGGAGAG 451  
DB 423 TGAAGATCATGAAGTGGACATGCTGAAGAGGCGAGGTGCTGCTTCCGTGAGG 482  
QY 452 AGCGGAACATATTATCTCGAAGACAAAGCCCGTGGATCCCGCAATACAGTATGCTTTTC 511  
DB 483 AGAGGACGCTTTGGTGAATGGGACCGCGGTGGATCAGCAGCTGCACTTCGCTTCC 542  
QY 512 AGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGGGACTTCTGTCAC 571  
DB 543 AGATGAGAACTACCTGCTACTGCTGATGAGTATACGTGGCGGGGACCTGCTGACAC 602  
QY 572 TTTTGAATAGATATGAGGACCAAGTATGATGAAACCTGATACAGTTTACCTAGCTGAGC 631  
DB 603 TGCTGAGCAAGTTTGGGAGCGGATTCGCGCCGAGATGGCGGCTTCTACCTGCGGAGA 662  
QY 632 TGATTTGGCTGTTCACAGCGTTTCATCTGATGGATGATGATGATGATGATGATGATG 691

Db 663 TTGTCATGCCCATAGACTCGTGCACCGGCTTGGCTAGCTGCACAGGACATCAAAACCG 722  
QY 692 AGAACAATCTGTTGACCCACAGACACATCAAGCTGTGTGATTTGGATCTGCCGGA 751  
Db 723 ACACATCTCTGTGACCGCTGTGGCCACATCGCTGGCCGACTTCGGCTTCGCTCA 782  
QY 752 AAATGAANTCAACAAGATGTTGAATGCAAACTCCGATTCGGACCCGACATACATGG 811  
Db 783 AGCTGGGCGAGATGGAACGGTGGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 842  
QY 812 CTCCTGAAGTGTGA---CTGTGATCAACGGGGATGGAAGGACACCTACGGCTGGACT 868  
Db 843 CCCCAGATCTCTGACGGCTGTGGGGGTGGGCTGGGACAGCAGCTACGGGCCGAGT 902  
QY 869 GTGACTGTGGTGGTGGGCTGTGCTTATGCTATGATGATTTATGGGAGATCCCCCTTCG 928  
Db 903 GTGACTGTGGGCTGTGGTGTATTCGCTATGAAATGTTCTATGGCAGACGCCCTTCT 962  
QY 929 CAGAGGGAACCTTGCAGAACCTTCAATAAATATGAATTTCCAGCGGTTTTGAAT 988  
Db 963 ACGGGATTCACGGCGGAGACTATGGCAAGATCGTCCACTACAGGAGACCTCTCTC 1022  
QY 989 TTCCAGATCACCCCAAGTGAGCAGT---GACTTTCTGATCTGATTCAAAGCTTGTGT 1045  
Db 1023 TGCCGCTGTGACGAGGGGTCCCTGAGGAGGCTCGAGACTTCATTACGGGTTGCTGT 1082  
QY 1046 GCGGCAGAAAAGAGACTGAAGTTTGAAGTCTTTGTGTC-----CATCCTTCT 1096  
Db 1083 GTCCCCGAGACACCGCTGGGCGGGTGGAGCAGGACTTCGGGACACATCCCTTCT 1142  
QY 1097 TCTTAAATTTGACTGGAACACATTCGTAATCTCTCTCCCTTCTGTTCCACCCCTCA 1156  
Db 1143 TCTTGGCTCTGACTGGGATGTCTCCGGGACAGGTCGCCCTTTACACCGGATTCG 1202  
QY 1157 AGTCTGACGATGACACTCCAAATTTGA 1184  
Db 1203 AAGTCCACCGACACATCAACTTGA 1230

RESULT 6

US-08-422-706B-12  
; Sequence 12, Application US/08422706B  
; Patent No. 5977333  
; GENERAL INFORMATION:  
; APPLICANT: Brook, J. David  
; APPLICANT: Housman, David E.  
; APPLICANT: Shaw, Duncan J.  
; APPLICANT: Harley, Helen G.  
; APPLICANT: Johnson, Keith J.  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC  
; TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02713  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/422,706B  
; FILING DATE: 14-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/284,543  
; FILING DATE: 08-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/023,612

; FILING DATE: 26-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/839,255  
; FILING DATE: 20-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/01545  
; FILING DATE: 19-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB93/00253  
; FILING DATE: 05-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB9202485.0  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: MIT-5830A2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2726 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-422-706B-12  
Query Match 3.4%; Score 220.8; DB 2; Length 2726;  
Best Local Similarity 53.7%; Pred. No. 6e-49;  
Matches 531; Conservative 0; Mismatches 442; Indels 15; Gaps 3;  
QY 212 CTGCTCTGATGAAGATTAAAGCAGTGAAGTTCGGAAGTTCGGAAGTATCCGACACCATAG 271  
Db 243 CCGAACTGGCCAGGACAAAGTACGTGGCCGACTTCTTCGAGTGGCGGAGCCCATCTGTGG 302  
QY 272 CTGAGTTACAGGAGCTCCAGCCTTCGGCAAGAGTTCGAAGTTCGAAGTTCGAGATTCTGAAGTGTATCGAC 331  
Db 303 TGAGGCTTAAGAGGTTCCGACTGCAGAGGAGCGACTTCGAGATTCTGAAGTGTATCGAC 362  
QY 332 GTGCTCACTTTGTGAAGTGCAGTGTGAAGAGAAAGCAACCGGGGACATCTATGCTA 391  
Db 363 CCGGGGCGTTTCAAGGAGTACGCGTGTGAAGATGAAGCAGACGCGGCGGAGTGTATGCCA 422  
QY 392 TGAAGTGTATGAAGAAGAGCTTTATTGGCCAGGAGGAGTTCATTTTTCAGGAAG 451  
Db 423 TGAAGATCATGAACAAGTGGGACATGCTGAAGAGGCGGAGGTGCTGCTTCGCTGAGG 482  
QY 452 AGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCGCAATTAAGTATGCTTTTC 511  
Db 483 AGAGGACGTTGTTGTTGAATGGGACCGGCGGTGGATCAAGCAGCTGCACTTCGCGCTTC 542  
QY 512 AGGACAAATACCTTTTATCTGATGAGGAATATCAGCCTGGAGGGAGCTTGTGTGAC 571  
Db 543 AGGATGAGAACTACCTGTACCTGTGATGAGTATTACGTTGGGCGGAGACTGCTGACAC 602  
QY 572 TTTTGAATAGATATCAGGACCCAGTTAGATGAAAACCTGATACAGTTTTCCTAGCTGAGC 631  
Db 603 TGCTGAGCAAGTTTGGGAGCGGATTCGGCGGAGATGGCGGCTTCTACCTGGCGGAGA 662  
QY 632 TGATTTGGCTGTTCACAGCGTTTCATCTGATGGATACGTCATCGAGACATCAAGCTG 691  
Db 663 TTGTCATGGCCATAGACTCGGTGACCGGCTTGGCTACGTCAGTGGGAGACATCAAAACCG 722  
QY 692 AGAATCTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTTGGATCTCCCGCGA 751  
Db 723 ACAATCTCTGCTGACCGCTGTGGCCACATCCGCTTGGCGGACTTCGGCTCTTGGCTCA 782  
QY 752 AAATGAATTTCAACAAGATGTTGAATGCCAAACTCCCGATTTGGGACCCCGAGATTACATGG 811  
Db 783 AGCTGGGCGAGATGGAACGGTGGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 842  
QY 812 CTCCTGAAGTGTGA---CTGTGATGAACGGGGATGGAAGGACACCTACGGCTGGACT 868

Db 843 CCCCCGAGATCCTCAGGCTGTGGCGGTGGCCCTGGGACAGGCGAGCTACGGGCGCGAGT 902  
Qy 869 GTGACTGTGTGTAGTGGCGGTGATTCCTATGAGATGATTTATGGGAGATCCCGCTTCG 928  
Db 903 GTGACTGTGTGGCGGTGATTCCTATGAGATGATTTATGGGAGATCCCGCTTCG 962  
Qy 929 CAGAGGAACTCTGCCAGAACCTTCAATAACATTAATGAATTTCCAGCGGTTCGAAAT 988  
Db 963 ACGCGGATTCACGGCGGAGACCTATGGCAAGATCGTCCACTACAGGAGCACTCTCTC 1022  
Qy 989 TTCAGATGACCCCAAGTGAAGT---GACITTTCTGTATCTGATTCATCAAGCTGTGT 1045  
Db 1023 TGGCGTGTGGAGCAAGGGTCCCTGAGGAGCTCGAGACTTCATTCAGCGGTGTGT 1082  
Qy 1046 GCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTTGTCTG-----CATCCTTTCT 1096  
Db 1083 GTCCCCCGGAGACACGGCTGGCGCGGTGGAGCAGGCGACTTCGCGACACATCCCTTCT 1142  
Qy 1097 TCTCTAAATGACTGGAACACATTCGTAACTCTCTCCCGCTTCGTTCCACCTCA 1156  
Db 1143 TCTTTGCGCTGAGTGGATGGTCTCCGGACAGCGTGGCCCCCTTTACACCGGATTCG 1202  
Qy 1157 AGTCTGAGATGACACCTCCAAATTTGA 1184  
Db 1203 AAGTGCCACGACACATGCACTTGA 1230

## RESULT 7

US-08-422-699A-8  
; Sequence 8, Application US/08422699A  
; Patent No. 5953265  
; GENERAL INFORMATION:  
; APPLICANT: Brook, J. David  
; APPLICANT: Housman, David E.  
; APPLICANT: Shaw, Duncan J.  
; APPLICANT: Harley, Helen G.  
; APPLICANT: Johnson, Keith J.  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC  
; TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02713  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/422,699A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/422,706  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/023,612  
; FILING DATE: 26-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/839,255  
; FILING DATE: 20-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/01545  
; FILING DATE: 19-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB93/00253  
; FILING DATE: 05-FEB-1993  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB9202485.0  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: MIT-5830A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2511 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1746  
US-08-422-699A-8  
Query Match 3.3%; Score 214; DB 2; Length 2511;  
Best Local Similarity 54.2%; Pred. No. 3.9e-47;  
Matches 508; Conservative 0; Mismatches 415; Indels 15; Gaps 3;  
Qy 262 GACACCATAGCTGAGTTACAGGAGCTCAGCCTTCGGCAAGAGACTTCGAGAGTCAGAGT 321  
Db 43 GCCATCGTGTGAGGCTTAAGGAGTCCGACTGCAGAGGAGCGACTTCGAGATTTCTGAAG 102  
Qy 322 CTGTGTAGTTGTGTCTCACTTTGCTGAAAGTGCAGGTGCTAAGAGAGAAAGCAACCGGGAC 381  
Db 103 GTGATCGACCGCGGGCGTTTCAGCGAGTAGCGGTAGTGAAGATGAAGAGACGCGGCAG 162  
Qy 382 ATCTATGCTATGAAGTGAAG 441  
Db 163 GTGTATGCCATGAAGATCATGAACAGTGGGACATGCTGAAGAGAGGCGGAGGTGCTGTC 222  
Qy 442 TTTGAGGAAGAGCGGAACATATATCTCGAAGCACACAGCCGCTGGATCCCGCATTAACAG 501  
Db 223 TCCGTGAGGAGAGGAGCGGTGTTGTTGAATGGGACCGCGGTGGATCAGCGAGCTCCAC 282  
Qy 502 TATCCCTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGGGAC 561  
Db 283 TTCGCTTCCAGGATGAGAACTACTCTGCTCATGAGTATTACGTGGCGGGGAC 342  
Qy 562 TTGCTGTCACTTTTGAATAGATATGAGGACAGTTAGATGAAACCTGATACAGTTTAC 621  
Db 343 CTGCTGACACTGTGAGCAAGTTTGGGAGGCGGATTCGCGCCGAGATGGCGGCTTCTAC 402  
Qy 622 CTAGCTGAGCTGATTTTGGCTTTTCAGCGGTTTCATCTGATGGGATACGTGCATTCGAGAC 681  
Db 403 CTGGCGGAGATTGTCATGGCCATAGACTCGTGCACCGGCTTGCTACGTGCACAGGAC 462  
Qy 682 ATCAAGCCTGAGAACATTCGTGTGACCGCACAGGACACATCAAGCTGGTGGATTTTGA 741  
Db 463 ATCAACCCGACACATCTCTGCTGGACCGCTGTGGCCACATCCGCTGGCGGCTTGGC 522  
Qy 742 TCTGCCGGAATAATGAATTCAAACAGATGTTGAATGCCAAACTCCCGATTTGGGACCCCA 801  
Db 523 TCTTGCTCAAGCTGCGGGCAGATGNAACGTTGCGGTGCTGCTGGTGGTGGTGGTGGT 582  
Qy 802 GATTACATGGCTCTGA---AGTGTGACTGTGATGAACGGGATGAAAGGACCTAC 858  
Db 583 GACTACCTGTCCCGGAGATCTCTGAGGCTGTGGCGGTGGCGCTGGGACAGGAGCTAC 642  
Qy 859 GGCTGAGCTGTGACTGTGTGCTGAGTGGGCTGATTCCTCTATCAGATGATTTATGGAGA 918  
Db 643 GGGCCGAGTGTGACTGTGTGGGCGCTGGGTGATTTGCTCTATGAATGTTTATGGGAG 702  
Qy 919 TCCCCCTTCGAGAGGAGACCTCTGCCAGAACCTTCAATAACATTAATGATTTCCAGCG 978  
Db 703 ACGCCCTTCTACGCGGATTCACGCGGAGACCTATGGCAAGATCGTCCACTACAGGAG 762

QY 979 TTTTGAATTTCCAGATGACCCCAAGTGAGCACT---GACTTCTTGATCTGATTCAA 1035  
Db 763 CACCTCTCTCTCCCGCTGGTGACGAGGGTCCCTGAGGAGGCTCGAGACTTCAATCAG 822  
QY 1036 AGCTTGTTGTGGGCGCAGAGAGAGACTGAAGTTTGAAGGCTTTTGTGTC----- 1086  
Db 823 CGTTGTGTGTCTCCCGGAGACGCTTGGCGGGGTGGAGCAGGCGACTTCCGGACA 882  
QY 1087 CATCTTTCTTCTTAAATTTGACTGGAAACAACTGTAATCTCTCTCCCGCTTCTGTT 1146  
Db 883 CATCCCTTCTTCTTGGCTCTGACTGGATGCTCTCCGGGACAGGCTCGCCCTTTACA 942  
QY 1147 CCCACCTCAAGTCTGACGATGACACCTCCCAATTTGA 1184  
Db 943 CCGGATTTGGAAGTCCACCGACACATGCAACTTCA 980

## RESULT 8

US-08-422-706B-8  
Sequence 8, Application US/08422706B  
Patent No. 5977333  
GENERAL INFORMATION:  
APPLICANT: Brook, J. David  
APPLICANT: Housman, David E.  
APPLICANT: Shaw, Duncan J.  
APPLICANT: Harlev, Helen G.  
APPLICANT: Johnson, Keith J.  
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC  
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02713  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,706B  
FILING DATE: 14-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/284,543  
FILING DATE: 08-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/023,612  
FILING DATE: 26-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/839,255  
FILING DATE: 20-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/01545  
FILING DATE: 19-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/00253  
FILING DATE: 05-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB9202485.0  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: MIT-5830A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:

; LENGTH: 2511 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1746  
; US-08-422-706B-8

Query Match 3.3%; Score 214; DB 2; Length 2511;  
Best Local Similarity 54.2%; Pred. No. 3.9e-47;  
Matches 508; Conservative 0; Mismatches 415; Indels 15; Gaps 3;

QY 262 GACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGSCAAAGGACTTCGAGTCAAGT 321  
Db 43 GCCATCGTGTGAGGCTTAAGAGGTCGACTCGAGAGGACGACTTCGAGATTCGAAG 102  
QY 322 CTTGTAGGTTTGTGTCACCTTTGTGAAGTCAGGTTGAAGAGAAAGCAACCGGGGAC 381  
Db 103 GTGATCGGACGCGGGGGCTTCAGCGAGGTAGCGGTAGTGAAGATGAAGCAGAGCGGGCAG 162  
QY 382 ATCTATCTATGAAGTGAATGATCAAGAGAGAGCTTTATTGGCCCGAGGAGGTTTCAATTT 441  
Db 163 GTGTATGCCATGAAGATCATGAACAAGTGGGACATGCTGAAGAGGGCGAGGTGCTGTC 222  
QY 442 TTTGAGGAAGAGCGGAAACATATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTACAG 501  
Db 223 TTCCGTGAGGAGGAGGACGTTTGGTGAATGGGACCGCGGTGGATCAACGAGCTGCAC 282  
QY 502 TATCCCTTTGAGGACAAATATCACCCTTTATCTGATGAGAGAAATATCAGCTTGGGGGAC 561  
Db 283 TTCGCTTCCAGGATGAAACTACCTCTACCTGCTCATGTAGTATTACGTGGGGGGGAC 342  
QY 562 TTGCTGTCACTTTTGAATAGATATGAGGACCGAGTATGATGAAACCTGATACAGTTTTAC 621  
Db 343 CTGCTGACATGCTGACCAAGTTTGGGAGGGGATTCGGCCCGAGATGGCGGCTTCTAC 402  
QY 622 CTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTTCAAGCGGATCATCTGATGGGATACGTGCA 681  
Db 403 CTGGCGGAGATTGTCTATGGCCATAGACTCGGTGACCGGCTTGGCTACGTGACAGGAC 462  
QY 682 ATCAAGCTGAGAACATCTCGTACCGGACACAGGACACATCAAGCTGTTGGATTTTGA 741  
Db 463 ATCAACCCGACAACTCCCTGCTGGACCGCTGTGGCCACATCCCTTGGCGACTTCGCG 522  
QY 742 TCTGCGCGAAATGAATTTCAACAAGATGTTGAATGCCAAACTCCCGATGGGACCCCA 801  
Db 523 TCTTGCCTCAAGCTCGGGCAGATGGAACGTCGCTCGCTGGTGGTGGGACCCCA 582  
QY 802 GATTACATGGCTCCTGA---AGTCTGACTGTGATGAACGGGGATGGAAGGACCTAC 858  
Db 583 GACTACCTGTCCCCCGAGATCCTGACGGCTGTGGGCGGTGGGCGCTGGGACAGGACGCTAC 642  
QY 859 GGCTTGACTGTGACTGTGCTGAGTGGGGCTGATTCCTATGATGATGATTTATGGGAGA 918  
Db 643 GGGCCCGAGTGTGACTGTGGGCGCTGGGTGATTCGCTATGAAATGTTCTATGGGAG 702  
QY 919 TCCCGCTTCGAGAGGGAACTCTGCGAGAACCTTCAATTAACATATGAATTTCCAGGG 978  
Db 703 ACGGCTTCTACGGCGATTCCACGGCGGAGACCTATGCGAAGATCGTCCACTACAAGAG 762  
QY 979 TTTTGAATTTCCAGATGACCCCAAGTGAAGT---GACTTCTTGATCTGATTCAA 1035  
Db 763 CACCTCTCTGCGCGCTGGTGGACGAAGGGTCCCTGAGGAGGCTCGAGACTTCATTCAG 822  
QY 1036 AGCTTGTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGTC----- 1086  
Db 823 CGGTGTGTGTCTCCCGGAGACACGCTTGGCGGGGTGGAGCAGGCGACTTCCGGACA 882  
QY 1087 CATCTTTCTCTCTAAATTTGACTGGAAACAACATTGTAATCTCTCTCCCGCTTCTGTT 1146  
Db 883 CATCCCTTCTTCTTGGCTCTGACTGGATGCTCTCCGGGACAGGCTGCGGCTTTTACA 942

QY 1147 CCCACCCCTCAAGTCTGACGATGACACCTTCCAAATTTGA 1184  
 Db 943 CCGGATTCGAGGTGCCACCGACACATGCAACTTGA 980

## RESULT 9

US-08-484-044-11  
 ; Sequence 11, Application US/08484044  
 ; Patent No. 5552282  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Caskey, C. T.  
 ; APPLICANT: Fu, Ying-Hui  
 ; APPLICANT: Friedman, David L.  
 ; APPLICANT: Pizzuti, Antonio  
 ; APPLICANT: Fenwick, Raymond G.  
 ; TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fulbright & Jaworski, L.L.P.  
 ; STREET: 1301 McKinney, Suite 5100  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: U.S.A.  
 ; ZIP: 77010-3095  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/484,044  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/019,940  
 ; FILING DATE: 19-FEB-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Paul, Thomas D.  
 ; REGISTRATION NUMBER: 32,714  
 ; REFERENCE/DOCKET NUMBER: D-5443  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 713/651-5325  
 ; TELEFAX: 713/651-5246  
 ; TELEX: 762829  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3182 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-08-484-044-11

Query Match 3.2%; Score 213.6; DB 1; Length 3182;  
 Best Local Similarity 54.2%; Pred. No. 5.8e-47;  
 Matches 507; Conservative 0; Mismatches 414; Indels 15; Gaps 3;  
 QY 264 CACCATAGCTAGTTACAGGAGCTCCAGCTTCGGCAAGGACTTCGAAGTCAGAGTCT 323  
 Db 743 CATCTGTGGTGGCTTAAGGAGGTCCGACTGCAGAGGGACGACTTCGAGATTCTGAAGGT 802  
 QY 324 TGTAGTTTGTGTCACCTTTGCTGAAGTCAGGTGGTGAAGAGAAAGCAACCGGGACAT 383  
 Db 803 GATCGACCGGGGCTTCAGCGAGGTAGCGGTAGTGAAGTGAAGCAGCGGGCCAGGT 862  
 QY 384 CTATGCTATCAAGTCATGAAGAGGCTTTATGGCCAGGAGAGGTTTCATTTT 443  
 Db 863 GTATGCTATCAAGTCATGAAGAGGCTTTATGGCCAGGAGAGGTTTCATTTT 922  
 QY 444 TGAGGAAGCGGAGCAATATTATCTGAAGCAACAGCCCGTGGATCCCCCAATTACAGTA 503  
 Db 923 CCTGTAGGAGAGGAGCGTGTGGTGAATCGGACCGCGGTGGATCAGCGAGCTGCACCT 982

QY 504 TGCCTTTTCCAGGACAAAATCACCTTTATCTGATGAGGAATATCAGCTTGAGGGGACTT 563  
 Db 983 CCGCTTTCCAGGATGAGAACTACCTGTACCTGGTTCATGGAGTATTACGTGGGCGGGGACCT 1042  
 QY 564 GCTGTCACTTTTGAATAGATATGAGGACCAAGTATAGATGAAAACCTGTATACAGTTTACCT 623  
 Db 1043 GCTGACACTGCTGACCAAGTTTGGGAGCGGATCCGGCCCGAGATGGCGCGCTTCTACT 1102  
 QY 624 AGCTGAGCTGATTTGGCTGTTTCACAGCGTTTCATCTGATGGGATACGTGCAATCCAGACAT 683  
 Db 1103 GCGCGAGATTGTCATGGCCATAGACTCGGTGCACCGGCTTGGCTACGTGTCACAGGGACAT 1162  
 QY 684 CAAGCTTGAGAAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGGATTTTGGATC 743  
 Db 1163 CAAGCCCGACACATCTGCTGGACCGCTGTGGCCACATCCGCTGGCCGACCTTCGGCTC 1222  
 QY 744 TGCCCGGAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACCCGAGA 803  
 Db 1223 TTGCTCAAGCTGCGGGCAGATGGAACGGTTCGGCTCGCTGTGGGACACCCGAGA 1282  
 QY 804 TTACATGCTCCTGA---AGTCTGACTGTGATGAACGGGATGGAAGGACACCTACGG 860  
 Db 1283 CTACTGTCCCCGAGATCTTCAGGCTGTGGCGGTGGGCTGGGACAGCAGCTACGG 1342  
 QY 861 CTTGACTGTGACTGGTGGTGGTGGGCTGATTGCTCTATGAGATGATTTATGGGAGATC 920  
 Db 1343 GCGGAGTGTGACTGGTGGGCTGGTGTATTGCTCTATGAAATGTTCTATGGGACAGC 1402  
 QY 921 CCCCTTCGACAGGAGACCTCTGCACAGACCTTCAATCAATATGAAATTCAGCGGTT 980  
 Db 1403 GCGCTTCTACGCGGATTCACGCGGAGACCTATGGCAAGATCGTCCACTACAGGAGCA 1462  
 QY 981 TTTGAAATTTCCAGATGACCCCAAGTGAGCAGT---GACTTTCTTGTATCTGATTTCAAAG 1037  
 Db 1463 CCTCTCTCGCGCTGGTGGAGGAGGGTCCCTGAGGAGGCTCGAGACTTCATTACGG 1522  
 QY 1038 CTTGTTGGCGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGTGTC-----CA 1088  
 Db 1523 GTTGTGTGTCCTCCCGGAGACACGGCTGGGCGGGGTGGAGCAGGCGACTTCGGGACACA 1582  
 QY 1089 TCCTTTCTTCTATAAATTCAGTGAACACATTCGTAACCTCTCTCCCTCCCTCCCTTCGTTCC 1148  
 Db 1583 TCCCTTTCTTTTGGCTGCTGAGTGGTCTCCGGGACAGCGTGTCCCTTTACACC 1642  
 QY 1149 CACCTCAAGTCTGACGATGACACCTCCAAATTTGA 1184  
 Db 1643 GGATTCGAAGGTGCCACCGACACATGCAACTTCGA 1678

## RESULT 10

US-09-804-471A-3  
 ; Sequence 3, Application US/09804471A  
 ; Patent No. 6479269  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WEBSTER, Marion et al  
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
 ; TITLE OF INVENTION: THEREOF  
 ; FILE REFERENCE: CL001164  
 ; CURRENT APPLICATION NUMBER: US/09/804,471A  
 ; CURRENT FILING DATE: 2001-03-13  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 174493  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)-(174493)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; US-09-804-471A-3



492 ATTTAAAGTTAGCTGACTTTTGGAACTGTATGAAATGGATACAGATGGTTGGTACGTT 551  
Db  
779 CCAAACTCCGATTTGGACCCAGATACATGCTGCTGAAAGTGGCTGACTGTGATGAAG 838  
Qy  
552 CTAATAATGCTGTGGAAACGCTGATACATTTCTCCCGAAGT-----TTTGCAGTCCC 605  
Db  
839 GGGATGAAAAGGACACCTACGGCTGGACTGTGACTGGTGTGCTCAGTGGCGCTGATTGCTT 898  
Qy  
506 AAGTGTGTGAAGGATTTACGGTGTGAATGCGATGGTGTCTGTGGGAATTTTGTGT 865  
Db  
899 ATGAGATGATTTATGGAGATGCCCCCTTCGCGAGGAACTCTGCCAGAACCTTCAATA 958  
Qy  
666 ATGAATGTTATTTGGAGAAACACCTTTTATGCGAGACAGTTTGGTGGAACTTACAGTA 725  
Db  
959 ACATTATGAATTTCCACGCGTTTGTGAAATTTCCAGATGACCCCAAGTGACAGTGACT 1018  
Qy  
726 AAATATGGATCACAGAACTCAATTAATTTTCTCCAGAGTGGAAATAAGCCAAATATG 785  
Db  
1019 TTCTTGATCTGATCAAGCTTTGTGGCGCCAGAAAGAG-----AGACTGAAGT 1069  
Qy  
786 CCGGATCTTTGATACAGGATTTTAAACAGACAGACACAGCGCTTTAGGCAGAAATGAAG 845  
Db  
1070 TTGAAGTCTTGTGCTGCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1123  
Qy  
846 TGAAGAAATTAACAGACATCCATTTTTCATAATGATCAATGGAGCTTTTGACAAATTA 905  
Db  
1124 GTAACCT 1183  
Qy  
906 GAGACTCTGCCCCACCTGTAGTCCAGAGCTGATGCTGATGATGATACAGGACTTTG 965  
Db  
1184 ATGA 1187  
Qy  
966 ATGA 969  
Db

## RESULT 13

US-09-005-069-61  
; Sequence 61, Application US/09005069  
; Patent No. 5932470  
; GENERAL INFORMATION:  
; APPLICANT: FRANK, GLENN R.  
; APPLICANT: HUNTER, SHIRLEY WU  
; APPLICANT: WALLENFELS, LYNDA  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS  
; TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shexidan Ross P.C.  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/005,069  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/630,822  
; FILING DATE: 11-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CONNELL, GARY J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-17-C3  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223

; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2706 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 5..2706  
; US-09-005-069-61

Query Match 2.9%; Score 189.2; DB 2; Length 2706;  
Best Local Similarity 54.1%; Pred. No. 2e-40;  
Matches 489; Conservative 0; Mismatches 388; Indels 27; Gaps 4;  
Qy 299 CAAGGACTTCGAAGTCAGAAAGCTCTGTAGTGTGTGCTACCTTTGCTGAAGTCGAGGTGG 358  
Db 78 CAGATGATTTTAAATTTAATAAAGTTATTTGTCGAGAGCATTGGTGAAGTACAGTTAG 137  
Qy 359 TAAGAGAGAAAGCAACCGGGGACATCTATGCTATGAAAGTGATGAAGAGCAAGCTTTAT 418  
Db 138 TCGGACACAAATCACTGCACAGTTTTCCTATGAAACGCTTATCAAAATTTGAATGA 197  
Qy 419 TGGCCCGAGGAGCAGGTTTCATTTTGGAGGAGCGGAAACATATATCTCTCGAAGCAAA 478  
Db 198 TTAAGAGACCCAGACTCTGCATTTTTCGGAAGAACCTCATATAATGGCTCATGCCAAA 257  
Qy 479 GCCCGTGGATCCCCCAATTACAGTATGCTTTCAGGACAAAATCACTTTATCTGATGG 538  
Db 258 CAGATGATGTACAAATTTATTTGCTTTTCAGATCAAAATATCTTTATATGCTCA 317  
Qy 539 AGGAATATCAAGCTGGAGGGAGCTTGTCTGCTCTTTTGAATAGATATAGGACCAAGTTAG 598  
Db 318 TGGATTATATGCCGGGGGTGACTTGGTGTAGTCTTATG-----TCGATTATGAAATTC 371  
Qy 599 ATGAAACCTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTTCACAGCGTTTCATC 658  
Db 372 CAGAAAATGGGCAATGTTCTTATACAAATGGAAGTGGTGTAGCATTGATCAATTCAC 431  
Qy 659 TGATGGGATACGTGCATCGAGACATCAAGCTCGAGAACATTCCTCGTTGACCGCACAGGAC 718  
Db 432 CCATGGGATTTGTACATCGTGTGTTAAACCTGATAATATGCTTCAGACAAATATGGTC 491  
Qy 719 ACATCAAGCTGGTGGATTTTGGATCTGCCCGGAAATGAATTCAAACAGATGGTGAATG 778  
Db 492 ATTTAAAGTTAGCTGACTTTTGGAACTCTGATGAAATGGATACAGATGGTTGCTGAGTT 551  
Qy 779 CCAAACTCCCGATTGGGACCCCGATTCATATGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 838  
Db 552 CTATAATGCTGTGGAAACGCTGATTCATTTCTCCGAGT-----TTTCAGTCCC 605  
Qy 839 GGGATGAAAAGGACACCTACGGCTGGACTGTGACTGTGTGCTGCTGCTGCTGCTGCTGCT 898  
Db 606 AAGTGTGTGAAGGAGTTTACGGTGTGAATTCGATTTGCTGCTGCTGCTGCTGCTGCTGCT 665  
Qy 899 ATGAGATGATTTATGGGAGATCCCTTCGACAGAGGAACTCTGCCAGAACCTTCAATA 958  
Db 666 ATGAATGTTATTTGGAGAAACACCTTTTATGCGAGACAGTTTGGTGGAACTTACAGTA 725  
Qy 959 ACATTATGAATTTCCAGCGGTTTGTGAAATTTCCAGATGACCCCAAGTGAGAGTGAAGT 1018  
Db 726 AAATATGGATCACAGAACTCAATTAATTTTCTCCAGAGTGGAAATAAGCCAAATATG 785  
Qy 1019 TTCTTGATCTGATCAAGCTTGTGTGGCGCCAGAAAGAG-----AGACTGAAGT 1069  
Db 786 CCGGATCTTTGATACAGGATTTTAAACAGACAGACACAGCGTTTAGGCAGAAATGAAG 845  
Qy 1070 TTGAAGTCTTGTGCTGCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1123  
Db 846 TGAAGAAATTAACAGACATCCATTTTTCATAATGATCAATGGAGCTTTTGACAAATTA 905  
Qy 1124 GTRACT 1183

Db 906 GAGACTCTGCCACCTGTAGTCCAGAGCTGAGTGGTATGATGATACAGAGAACTTTG 965  
QY 1184 ATGA 1187  
Db 966 ATGA 969

RESULT 14  
US-09-171-156A-20  
; Sequence 20, Application US/09171156A  
; Patent No. 6368846  
; GENERAL INFORMATION:  
; APPLICANT: Hunter, Shirley Wu  
; Sim, Gek-Ke  
; Weber, Eric R.  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND  
; APPARATUS TO COLLECT SUCH PROTEINS  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS P.C.  
; STREET: 1560 BROADWAY, SUITE 1200  
; CITY: DENVER  
; STATE: CO  
; COUNTRY: U.S.A.  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/171,156A  
; FILING DATE: 04-Mar-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2706 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 5..2706  
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-171-156A-20  
Query Match 2.9%; Score 189.2; DB 4; Length 2706;  
Best Local Similarity 54.1%; Pred.No.2e-40;  
Matches 489; Conservative 0; Mismatches 388; Indels 27; Gaps 4;  
QY 299 CAAAGACTCGAAGTCAGAACTCTGTAGTGTGGTCACTTTGCTGCAAGTCGAGGTGG 358  
Db 78 CAGATGATTTAAATTAATAAAGTTATTGGTCGAGGAGCACTTTGGTCAAGTCACAGTTAG 137  
QY 359 TAAGAGAGAACACCGGGACATCTATGCTATGAAGTATGAGAGAGGCTTTAT 418  
Db 138 TCGCACACAAATCAACTGCACAAAGTTTGTCTATGAACGGCTATCAAAATTTGAATGA 197  
QY 419 TGGCCCGAGGACAGGTTTCATTTTTTGGAGGAGCGGAACATATATCTCGAAGCAAA 478  
Db 198 TTAAGAGACCACTCTGCAATTTTTTGGAGAGACGTCATATATGCTCATCAAAAT 257  
QY 479 GCCCGTGAATCCCCCAATACAGTATGCTTTCAGACAAAAATCAACCTTTATCTATGATGG 538

Db 258 CAGAATGGATTGTACAAATTACATTTTCTCTTTCAAGATCAAAATATCTTTATATGTC 317  
QY 539 AGGAATATACGCTCGAGGGAGCTTGTCTGTCACTTTTGAATAGATATGAGGACCAAGTTAG 598  
Db 318 TGGATTATATGCCGGGGGTGACTTGGTGAATCTTATG-----TCCGATTATCAAAATTC 371  
QY 599 ATGAAAACCTGTACACAGTTTTTACCTAGCTGAGTGTATTTTGGGTGTTTCCAGGGTTTCATC 658  
Db 372 CAGAAAATATGGCAATGTTCTATACATGGAAGTGGTGTAGCACTTGATACATTCAC 431  
QY 659 TGATGGATACGTGCATCGAGACATCAAGCTCGAGAACATCTCTGTTGACCGCACAGGAC 718  
Db 432 CCATGGGATTTGTACATCGTGTATTAACCTGATATATCTTCTAGACAAATATGGTC 491  
QY 719 ACATCAGCTGGTGGATTTTGGATCTGCCCGCAAAATGATTCAAAACAAATCGTGAATG 778  
Db 492 ATTTAAAGTTAGTGTGCTTTTGGAACTGTATGAAATGGATACAGATGGTTTGTAGCTT 551  
QY 779 CCAAACTCCCGATTGGGACCCAGATTACATGCTCTCTGAAGTGTCTGACTGTGATGAACG 838  
Db 552 CTAATAATGCTGTTGGAACGCTGATTACATTTCTCCGGAAGT-----TTTGCAGTCCC 605  
QY 839 GGGATGGAAGGACCTAGCGCTGAGCTGTGACTGTGCTGCTGCTGCTGCTGCTGCTGCT 898  
Db 606 AAGTGGTGAAGGAGTTTACGGTCTGATGCGAATGGTGGTCTGTGGGAATTTTTTTGT 655  
QY 899 ATGAGATGATTTATGGAGATCCCTCTCGCAGAGGGAACCTCTGCCAGAACCTTCAATA 958  
Db 666 ATGAATGTTATTTGGAGAACACCTTTTATGCAGACAGTTTGGTGGAACTTACAGTA 725  
QY 959 ACATTATGAATTTCCAGCGGTTTTTGAATTTCCAGATGACCCCAAGTGTGAGCAGTGACT 1018  
Db 726 AAATTATGATCACAGAACTCATTAACCTTCTCCAGAAAGTGGAAATGAAGCAATATG 785  
QY 1019 TTCTTGATCTGATTCAAAGCTGTGTGGCGGCCAGAAAGAG-----AGACTGAAGT 1069  
Db 786 CCCGATCTTGATACAGAGATTTTAAACAGACAGACAGACAGCGTTTGGCAGAAATGAAG 845  
QY 1070 TTGAAGTCTTTGCTGCCATCTTTCTCTCTAAAATTTGCTGGA-----ACAAATTC 1123  
Db 846 TGAAGAAATTAACAGACATCCATTTTTCATAATGATCAATGAGCTTTTGAACAATTTAA 905  
QY 1124 GTAACTCTCTCCCTCCCTTCTGTTCCACCTCAAGTCTGAGGATGACACCTCCCAATTTG 1183  
Db 906 GAGACTCTGCCACCTGTAGTCCAGAGCTGAGTGGTATGATGATACAGGAACTTTG 965  
QY 1184 ATGA 1187  
Db 966 ATGA 969

RESULT 15  
US-09-004-730A-20  
; Sequence 20, Application US/09004730A  
; Patent No. 6485968  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eric  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Sim, Gek-Ke  
; APPLICANT: Frank, Glenn  
; APPLICANT: Wallenfels, Lynda  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO COLLECT SUCH  
; FILE REFERENCE: 2618-17-C5-PUS-1  
; CURRENT APPLICATION NUMBER: US/09/004,730A  
; CURRENT FILING DATE: 1998-01-08  
; PRIOR APPLICATION NUMBER: PCT/97US/18669  
; PRIOR FILING DATE: 1997-10-15  
; NUMBER OF SEQ ID NOS: 150  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 2706  
; TYPE: DNA  
; ORGANISM: Ctenocephalides felis

Search completed: July 4, 2004, 09:51:43  
Job time : 402 secs

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(2704)
; US-09-004-730A-20

Query Match      2.9%; Score 189.2; DB 4; Length 2706;
Best Local Similarity 54.1%; Pred. No. 2e-40;
Matches 489; Conservative 0; Mismatches 388; Indels 27; Gaps 4;

Qy 299 CAAAGGACTTCGAAAGTCAGAAAGTCCTGAGTGTGTGTCACATCTTTGCTGAAGTGCAGGTGG 358
Db CAGATGATTTTAAATTTAATAAAGTTATTTGGTCGAGGAGCATTTGGTGAAGTACAGTTAG 137
Qy 359 TAAGAGAGAAAGCAACCGGGACATCTATGCTATGAAAGTGATCAAGAAAGCGCTTTAT 418
Db TGGGACACAATCAACTGCACAAGTTTTCGTATGAAACCGCTATCAAAATTTGAAATGA 197
Qy 419 TGGCCCGAGGAGCGGTTTCATTTTTCGAGGAAGCGGAACATATTTATCTCGAAGCACAA 478
Db TTAAGAGACAGACTCTGCACTTTTTCGGAAGAAACGTCATATAATGGCTCATGCCAAAT 257
Qy 479 GCCCGTGATCCCCCAATTACAGTATGCGCTTCAGGACAAAATTCACCTTTATCTGATGG 538
Db CAGATGGATTGTACAATTACATTTTGCTTTTCAAGATCAAAATATCTTTTATAGTCA 317
Qy 539 AGGAATATCAGCCTGGAGGAGCTGTGTCACCTTTTGAATAGATATGAGGACGAGTTAG 598
Db TCGATTATATGCCGGGGGCTGACTTGGTGAAGTCTATG-----TCCGATTATCAAAATTC 371
Qy 599 ATGAAACCTGTATACAGTTTTCCTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTCACT 658
Db CAGAAATATGGGCATGTTCTTACAAATGGAAGTGGTCTAGCACTTGTATCAAAATTCCT 431
Qy 659 TGATGGATACGTGTCATCGAGACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGAC 718
Db CCATGGGATTGTACATCGTGATGTTAAACCTGATAATATGCTTCTAGACAAATATGGTC 491
Qy 719 ACATCAAGCTGTGGATTTTGGATCTCCCGGAAATGAATTTCAACAAGATGGTGAATG 778
Db ATTTAAAGTTAGCTGACTTTTGGAACTGTATGAAATGGATACAGATGGTTTGGTAGCTT 551
Qy 779 CCAAACTCCCGATTGGGACCCAGATTACATGGCTCCTGAACTGCTGACTGTGATGAACG 838
Db CTAATAATGCTGTTGGAAACGCTGATTACATTTCTCCCGAAGT-----TTGCAGTCCC 605
Qy 839 GGAATGAAAGGACCTACGGCTGCACTGTGACTGTGTCAGTGGTGGTGGGCGTGAATGCT 898
Db AAGGTGCTGAAGGAGTTTACCGTCTGGAATGCGAATGGCTGCTGTGGGGAATTTTGTGT 665
Qy 899 ATGAGATGATTTATGGGAGATCCCGCTTCGACAGGGAACTCTGCGCAGAACCTTCAATA 958
Db ATGAAATGTTATTGGAGAACACCTTTTATGCAGACAGTTTGGTTGGAACATTACAGTA 725
Qy 959 ACATATGAAATTTCCACGGGTTTGGAAATTTCCAGATGACCCCAAGTGCAGAGTGACT 1018
Db AAATATATGGATCAGAAACTCATTAATCTTTTCTCCAGAAAGTGGAAATTAAGCCAAATATG 785
Qy 1019 TTCTTGATCTGATTCAAAAGCTTGTGTGCGGCGCAGAAAGAG-----AGACTGAAGT 1069
Db CCGATCTTTGATACAGGATTTTAAACAGACAGAACACACGCTTTAGGCAGAAATGAAG 845
Qy 1070 TTGAAGTCTTTTGTGTCATCTCTTTCTCTTAAATTGACTGGA-----ACAACATTC 1123
Db TGGAAAGAAATTAACAGCATCCATCTTTTTCATAATGATCAATGGACTTTTGAACAATTTAA 905
Qy 1124 GTAACCTCTCTCCCGCTTCTGTTCCCACTCCAGTCTGACGATGACACCTCCAAATTTTG 1183
Db GAGACTCTGCCCACTGTAGTCCAGAGCTGAGTGGTGTATGATGATCAAGGACCTTGTG 965
Qy 1184 ATGA 1187
Db ATGA 969

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 18:44:52 ; Search time 2174 Seconds  
(without alignments)  
12846.207 Million cell updates/sec

Title: US-10-017-216-1  
Perfect score: 6574  
Sequence: 1 agagccgacagtgaggagat.....atccagaatgaggtttaga 6574

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6574	100.0	6574	6 AAD39191	Aad39191 Human MDP
2	5684.4	86.5	6298	6 AAD38864	Aad38864 Human kin
3	5668	86.2	8603	8 AAL55215	Aal55215 Human CRI
4	5666.6	86.2	6156	8 AAL55217	Aal55217 Human CRI
5	5666.2	86.2	6165	8 AAL55214	Aal55214 Human CRI
6	5661.4	86.1	6165	6 ABQ78870	Abq78870 Human kin
7	5650.4	86.0	6159	4 AAS06701	Aas06701 Polynucle
8	5631	85.7	6189	6 ABS63436	Abs63436 RHO/RAC-1
9	5631	85.7	6189	7 ADA05641	Ada05641 Human NOV
10	5629.8	85.6	6201	6 ABS63435	Abs63435 Human cDN
11	5629.8	85.6	6201	7 ADA05653	Ada05653 Human NOV
12	5358	81.5	5877	6 ABQ78871	Abq78871 Human kin
13	3475.2	52.9	6809	3 AAC77568	Aac77568 Human ORF
14	2575.2	39.2	2896	5 ABV30132	Abv30132 Human pro
15	2436.8	39.0	3131	4 ABA08361	Aba08361 Human RHO
16	2436.8	37.1	5261	8 AAL55216	Aal55216 Human CRI
17	2426.8	36.9	5251	9 ADD89966	Add89966 Human can
18	2414	36.7	2542	7 ADA05647	Ada05647 Human NOV
19	2317.2	35.2	2497	7 ADA05645	Ada05645 Human NOV
20	2126.8	32.4	2893	9 ADS09823	Ads09823 Novel DNA
21	1841.2	28.0	1870	7 ADA05643	Ada05643 Human NOV
22	1524.2	23.2	1870	7 ADA05649	Ada05649 Human NOV
23	1467.6	22.3	1915	7 ADA05651	Ada05651 Human NOV

ALIGNMENTS

RESULT 1									
AAD39191									
ID	AAD39191	standard;	cdna;	6574	BP.	24	1406.4	21.4	2066
XX						25	1397.8	21.3	1515
AC	AAD39191;					26	1397.8	21.3	1515
XX						27	1395	21.2	1485
DT	04-OCT-2002	(first entry)				28	1395	21.2	1765
XX						29	1087.8	16.5	2380
DE	Human MDPK CDNA.					30	965.6	14.7	2162
XX						31	943.4	14.4	1048
KW	Human; myotonic dystrophy type protein kinase; MDPK; 13245 protein;					32	701.4	10.7	995
KW	tumorigenesis; tumour growth; tumour metastasis; viral infection;					33	648	9.9	1058
KW	skeletal muscle disorder; muscular dystrophy; myotonic dystrophy;					34	592	9.0	817
KW	immune disorder; neoplastic disorder; gene therapy; gene; ss.					35	443.4	6.7	446
XX						36	436	6.6	485
OS	Homo sapiens.					37	433.2	6.6	580
XX						38	373.2	5.7	398
FH	Key	Location/Qualifiers				39	350.4	5.3	354
FT	5'UTR	1..18				40	337	5.1	396
FT		/*tag= a				41	308	4.7	308
FT	CDS	19..6180				42	306.4	4.7	308
FT		/*tag= b				43	306.4	4.7	308
FT		/product= "Human MDPK protein"				44	258	3.9	258
FT	misc_feature	19..6177				45	250	3.8	254
FT		/*tag= c							
FT		/note= "This region is specifically referred in claim 1							
FT		as SEQ ID NO:3"							
FT	3'UTR	6181..6574							
FT		/*tag= d							
XX									
FN	WO200234896-A2.								
XX									
PD	02-MAY-2002.								
XX									
PF	23-OCT-2001; 2001WO-US050636.								
XX									
PR	23-OCT-2000; 2000US-0242429P.								
XX									
PA	(MILL-) MILLENNIUM PHARM INC.								
XX									
PI	Kapeller-Libermann R;								
XX									
DR	WPI; 2002-479720/51.								
XX									
DR	P-PSDE; AAE24079.								
XX									
PT	Human myotonic dystrophy type protein kinase polypeptide and								

PT polynucleotide useful for prognosticating, diagnosing, preventing or  
PT inhibiting tumorigenesis, tumor growth, tumor metastasis and viral  
XX infection.

PS Claim 1; Fig 1; 148pp; English.

XX The invention relates to human myotonic dystrophy type protein kinase  
CC (MDPK) polypeptides designated as 13245 and nucleic acid molecules  
CC encoding such polypeptides. 13245 molecules are used to develop  
CC diagnostic and therapeutic agents for prognosticating, diagnosing,  
CC preventing, inhibiting, alleviating or curing MDPK-related disorders.  
CC Polypeptides of the invention are used to develop diagnostic and  
CC therapeutic agents for 13245-mediated or related disorders such as  
CC tumorigenesis, tumour growth, tumour metastasis, viral infection of a  
CC cell, skeletal muscle disorders (e.g. muscular and myotonic dystrophies),  
CC immune disorders and neoplastic disorders. The invention is also used in  
CC gene therapy. The present sequence is human MDPK cDNA

XX Sequence 6574 BP; 1877 A; 1611 C; 1776 G; 1310 T; 0 U; 0 Other;

Query Match 100.0%; Score 6574; DB 6; Length 6574;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6574; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGAGCCGCCAGTGGGAGATGTTGAAGTTCAATATGGAGCGCGAATCCTTTGGATGCT	60
DB	1	AGAGCCGCCAGTGGGAGATGTTGAAGTTCAATATGGAGCGCGAATCCTTTGGATGCT	60
QY	61	GGTGCTGCTGAACCCATTCGACCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGAAA	120
DB	61	GGTGCTGCTGAACCCATTCGACCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGAAA	120
QY	121	CCACCTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAGGGATATTAGATGC	180
DB	121	CCACCTTTATGACTCAACAGCAGATGTCCTCTTTCCGAGAGGGATATTAGATGC	180
QY	181	CTCTTTGTTCTTTGAAGATCAGTCAGCCTGCTCTGATGAAGATTAGCAGTGAGC	240
DB	181	CTCTTTGTTCTTTGAAGATCAGTCAGCCTGCTCTGATGAAGATTAGCAGTGAGC	240
QY	241	AATTTGTCGGAAGATTCCGACACATAGCTAGTTACAGAGCTCCAGCCTTCGGCA	300
DB	241	AATTTGTCGGAAGATTCCGACACATAGCTAGTTACAGAGCTCCAGCCTTCGGCA	300
QY	301	AAGGACTTCGAAGTCAGAGTCTTTAGGTTGTTGTTCACTTTGCTGAAGTCAGTGGTA	360
DB	301	AAGGACTTCGAAGTCAGAGTCTTTAGGTTGTTGTTCACTTTGCTGAAGTCAGTGGTA	360
QY	361	AGAGAGAACCAACCGGGACATCTATGATGAAGTATGAGNAGAGGCTTTATTG	420
DB	361	AGAGAGAACCAACCGGGACATCTATGATGAAGTATGAGNAGAGGCTTTATTG	420
QY	421	GCCAGGAGCAGGTTTCATTTTGAAGAGAGCGGAACATATTATCTCAAGCAACAGC	480
DB	421	GCCAGGAGCAGGTTTCATTTTGAAGAGAGCGGAACATATTATCTCAAGCAACAGC	480
QY	481	CCGTGATCCCCCAATTACAGTATGCTTTTCAAGACAAAATCACCCTTTATCTGATGGAG	540
DB	481	CCGTGATCCCCCAATTACAGTATGCTTTTCAAGACAAAATCACCCTTTATCTGATGGAG	540
QY	541	GAATATCAGCTCGAGGGGACTTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGAT	600
DB	541	GAATATCAGCTCGAGGGGACTTGTCTCACTTTTGAATAGATATGAGGACCAAGTTAGAT	600
QY	601	GAATACTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTGATCTG	660
DB	601	GAATACTGATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCGTTGATCTG	660
QY	661	ATGGGATACGTGATCGAGACATCAAGCTTGAGAACATTTCTGTTGACCGCAGAGACAC	720
DB	661	ATGGGATACGTGATCGAGACATCAAGCTTGAGAACATTTCTGTTGACCGCAGAGACAC	720
QY	721	ATCAAGCTGGTGGATTTTGGATCTGCCGCAAAATGAATTCAAACAGATGTTGATGCC	780
DB	721	ATCAAGCTGGTGGATTTTGGATCTGCCGCAAAATGAATTCAAACAGATGTTGATGCC	780

DB	721	ATCAAGCTGGTGGATTTTGGATCTGCCGCAAAATGAATTCAAACAGATGTTGATGCC	780
QY	781	AAACTCCCGATTGGACCCAGATATACATGGCTCTCTGAAGTGTGATGAACGGG	840
DB	781	AAACTCCCGATTGGACCCAGATATACATGGCTCTCTGAAGTGTGATGAACGGG	840
QY	841	GATGAAAGGCCACTACGCGCTGGATCTGCTGGTCTCACTGGGCGTGATGGCTAT	900
DB	841	GATGAAAGGCCACTACGCGCTGGATCTGCTGGTCTCACTGGGCGTGATGGCTAT	900
QY	901	GAGATGATTTATGGAGATCCCTCTTCGACAGAGGAACCTCTCCAGAACCTTCAATAAC	960
DB	901	GAGATGATTTATGGAGATCCCTCTTCGACAGAGGAACCTCTCCAGAACCTTCAATAAC	960
QY	961	ATTATGAAATTCAGCGGTTTGTGAATTTCCAGATGACCCCAAGTGACGACTGCTTT	1020
DB	961	ATTATGAAATTCAGCGGTTTGTGAATTTCCAGATGACCCCAAGTGACGACTGCTTT	1020
QY	1021	CTTGATCTGATTCAAAGCTTGTGTGGCGCCAGAAAGAGAGACTGAAGTTTGAAGTCTT	1080
DB	1021	CTTGATCTGATTCAAAGCTTGTGTGGCGCCAGAAAGAGAGACTGAAGTTTGAAGTCTT	1080
QY	1081	TGCTGCCATCTTTCTCTCTAAAATTGACTGGAACAACATTCGTAACCTCTCTCCCTCC	1140
DB	1081	TGCTGCCATCTTTCTCTCTAAAATTGACTGGAACAACATTCGTAACCTCTCTCCCTCC	1140
QY	1141	TTCCGTTCCCACTCAAGTCTGACGATGACACTCCAAATTTGATGAACCCAGAGAGAAAT	1200
DB	1141	TTCCGTTCCCACTCAAGTCTGACGATGACACTCCAAATTTGATGAACCCAGAGAGAAAT	1200
QY	1201	TCGTGGTTCATCTCTCCGTGACGCTGAGCCCTCAGGCTTCCTGGGTGAAGAACTG	1260
DB	1201	TCGTGGTTCATCTCTCCGTGACGCTGAGCCCTCAGGCTTCCTGGGTGAAGAACTG	1260
QY	1261	CCGTTTGGGGTTTCGTACAGCAGGCACTGGGATCTCTGTAGATCTGATCTGTT	1320
DB	1261	CCGTTTGGGGTTTCGTACAGCAGGCACTGGGATCTCTGTAGATCTGATCTGTT	1320
QY	1321	GTGTGGGTTTCGAACTCCCTGCAAGACTAGCTCCATGGAAGAAACTTCTCATCAA	1380
DB	1321	GTGTGGGTTTCGAACTCCCTGCAAGACTAGCTCCATGGAAGAAACTTCTCATCAA	1380
QY	1381	AGCAAGAGCTACAGACTCTCAGGACAGTGTCAAGATGAGCAGGAGAAATGACCCGG	1440
DB	1381	AGCAAGAGCTACAGACTCTCAGGACAGTGTCAAGATGAGCAGGAGAAATGACCCGG	1440
QY	1441	TTACATCGGAGTGTGAGGCTGCTGTAGTCAGAGGAGGTTGAGGCTGAAG	1500
DB	1441	TTACATCGGAGTGTGAGGCTGCTGTAGTCAGAGGAGGTTGAGGCTGAAG	1500
QY	1501	GCCTCTCAGACTCAGAGTCCCTCTCGGACAGGACCTTCTACCTACATCAGAAATGC	1560
DB	1501	GCCTCTCAGACTCAGAGTCCCTCTCGGACAGGACCTTCTACCTACATCAGAAATGC	1560
QY	1561	AGTAGCTTAAAGCGAAGTTTGGAGCAAGCAGGATGGAGTGTCCCAAGAGGATGACAAA	1620
DB	1561	AGTAGCTTAAAGCGAAGTTTGGAGCAAGCAGGATGGAGTGTCCCAAGAGGATGACAAA	1620
QY	1621	GCACTGACGCTTCTCCATGATATCAGAGACAGAGCGGAGGCTCCAAAGAAATCAAAGAG	1680
DB	1621	GCACTGACGCTTCTCCATGATATCAGAGACAGAGCGGAGGCTCCAAAGAAATCAAAGAG	1680
QY	1681	CAGAGTACCAAGCTCAAGTGGAGAAATGAGTGTGATGATGATCACTTGGAGAGGAT	1740
DB	1681	CAGAGTACCAAGCTCAAGTGGAGAAATGAGTGTGATGATGATCACTTGGAGAGGAT	1740
QY	1741	CTTCTCTCAGCAAGAGACGAGTGTCTCTACGAATCTCAGCTGAGAGAGTCTCGGCTT	1800
DB	1741	CTTCTCTCAGCAAGAGACGAGTGTCTCTACGAATCTCAGCTGAGAGAGTCTCGGCTT	1800
QY	1801	GCTCTGAGAAATTCAGCGGAAGCCAGCAATGTGACGATTAACCTGTTGAAGGCTAAG	1860
DB	1801	GCTCTGAGAAATTCAGCGGAAGCCAGCAATGTGACGATTAACCTGTTGAAGGCTAAG	1860

Db 1801 GCTGCTGAAGAAATTCAGCGGAAAGCGACAGAAATGTACGATATAAACTGTTGAAGCTAAG 1860  
Qy 1861 GATCAAGGGAAGCCTCAAGCTGGGAGAAATATCGAAACTCGAGAAGATCAATCTCAGCAG 1920  
Db 1861 GATCAAGGGAAGCCTCAAGCTGGGAGAAATATCGAAACTCGAGAAGATCAATCTCAGCAG 1920  
Qy 1921 CAGCTCAAAATTCAGGAGCTCCAAGAGAAACTGGAGAGCTGCAAGAGCGAGCCGAG 1980  
Db 1921 CAGCTCAAAATTCAGGAGCTCCAAGAGAAACTGGAGAGCTGCAAGAGCGAGCCGAG 1980  
Qy 1981 AGGAGCTCGAGAAGCTCGAGAACCGCGCCATCTCTGGAGAACAAAGTAAAGAGACTAGACCATG 2040  
Db 1981 AGGAGCTCGAGAAGCTCGAGAACCGCGCCATCTCTGGAGAACAAAGTAAAGAGACTAGACCATG 2040  
Qy 2041 GTGGAAGCTGAGGAACCGCGCCATCTCTGGAGAACAAAGTAAAGAGACTAGACCATG 2100  
Db 2041 GTGGAAGCTGAGGAACCGCGCCATCTCTGGAGAACAAAGTAAAGAGACTAGACCATG 2100  
Qy 2101 GAGCGTAGAGAAAACAGACTGAAGTATCATCTCAGAACAAATCCCAAGATCCAGCAG 2160  
Db 2101 GAGCGTAGAGAAAACAGACTGAAGTATCATCTCAGAACAAATCCCAAGATCCAGCAG 2160  
Qy 2161 ATGGCTGATAAAATTCGAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTCAGCCCGAG 2220  
Db 2161 ATGGCTGATAAAATTCGAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTCAGCCCGAG 2220  
Qy 2221 CACCTAGAGTGCACCTGAAAACAGAAAGAGAGCACTATAGGAAAAGATTAAAGTGTG 2280  
Db 2221 CACCTAGAGTGCACCTGAAAACAGAAAGAGAGCACTATAGGAAAAGATTAAAGTGTG 2280  
Qy 2281 GACAAATCAGATAAAGAAAGCACTGGCTGACAAAGGAGACACTGAGAAACATGATGACAGAGA 2340  
Db 2281 GACAAATCAGATAAAGAAAGCACTGGCTGACAAAGGAGACACTGAGAAACATGATGACAGAGA 2340  
Qy 2341 CACGAGGAGGCGCCATCAGAAAGGCAAAATTTCTCAGCGAAACAGAAAGCGGATGATCAAT 2400  
Db 2341 CACGAGGAGGCGCCATCAGAAAGGCAAAATTTCTCAGCGAAACAGAAAGCGGATGATCAAT 2400  
Qy 2401 GCTATGATTCAGAAATCAGATCCCTGGAACAGAGAGATTGTGAATCTGTGAAGCCAAAT 2460  
Db 2401 GCTATGATTCAGAAATCAGATCCCTGGAACAGAGAGATTGTGAATCTGTGTGAAGCCAAAT 2460  
Qy 2461 AAACCTTGACGAAATAGCAGTCTTTTACCCAAAGAAACATGAAGGCCCAAGAGAGATG 2520  
Db 2461 AAACCTTGACGAAATAGCAGTCTTTTACCCAAAGAAACATGAAGGCCCAAGAGAGATG 2520  
Qy 2521 ATTTCTGAATCAGGCAACAGAAAATTTTACCTGGAGACACAGGCTGGAAAGTTGAGGCC 2580  
Db 2521 ATTTCTGAATCAGGCAACAGAAAATTTTACCTGGAGACACAGGCTGGAAAGTTGAGGCC 2580  
Qy 2581 CAGAACCGGAAACTGGAGGAGCAGCTGGAGAGATCAGCCACCAAGACACAGAGCCACAG 2640  
Db 2581 CAGAACCGGAAACTGGAGGAGCAGCTGGAGAGATCAGCCACCAAGACACAGAGCCACAG 2640  
Qy 2641 AATCGGCTGCTGGAATCGGAGACAGATTGCGGAGGTCAGTCTAGACACGAGAGCAG 2700  
Db 2641 AATCGGCTGCTGGAATCGGAGACAGATTGCGGAGGTCAGTCTAGACACGAGAGCAG 2700  
Qy 2701 AACTGGAGCTCAAGGCCAGCTCAGAGCTACAGCTCTCCTCGAGAGCGGAGTCA 2760  
Db 2701 AACTGGAGCTCAAGGCCAGCTCAGAGCTACAGCTCTCCTCGAGAGCGGAGTCA 2760  
Qy 2761 CAGTTGACAGCCCTGAGGCTGACGGGGCGCTGGAGAGCAGCTTCGCCAGCGGAG 2820  
Db 2761 CAGTTGACAGCCCTGAGGCTGACGGGGCGCTGGAGAGCAGCTTCGCCAGCGGAG 2820  
Qy 2821 ACAGAGCTGGAAGAGACCAACAGAGCTGAAGAGAGATCCAGGCACTCAGGCAAT 2880  
Db 2821 ACAGAGCTGGAAGAGACCAACAGAGCTGAAGAGAGATCCAGGCACTCAGGCAAT 2880  
Qy 2881 AGAGATGAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAATCAGAC 2940  
Db 2881 AGAGATGAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAATCAGAC 2940

Qy 2941 CTGAGGAGCAGCTAAACCAGCTGACCGAGCAACCGTGAATCTCAACAAACCAAACTTC 3000  
Db 2941 CTGAGGAGCAGCTAAACCAGCTGACCGAGCAACCGTGAATCTCAACAAACCAAACTTC 3000  
Qy 3001 TACTTGTCCAAACAACTCGATGAGGCTTCTGGGCCCAACGACGAGATTGTACAACTCGA 3060  
Db 3001 TACTTGTCCAAACAACTCGATGAGGCTTCTGGGCCCAACGACGAGATTGTACAACTCGA 3060  
Qy 3061 AGTCAAGTGAACCACTCTCCCGGGAGATCACGGAACGAGAGATGACGCTTACCAGCAG 3120  
Db 3061 AGTCAAGTGAACCACTCTCCCGGGAGATCACGGAACGAGAGATGACGCTTACCAGCAG 3120  
Qy 3121 AAGCAAAACGATGAGGCTCTGAAGACCACTGACCACTGTGGAGGAAACAGGTCTATGAT 3180  
Db 3121 AAGCAAAACGATGAGGCTCTGAAGACCACTGACCACTGTGGAGGAAACAGGTCTATGAT 3180  
Qy 3181 TTGAGAGCCCTAAACGATGAGCTCTAGAAAAGAGCGGAGTGGAGGCTGGAGAGC 3240  
Db 3181 TTGAGAGCCCTAAACGATGAGCTCTAGAAAAGAGCGGAGTGGAGGCTGGAGAGC 3240  
Qy 3241 GTCCTGGTGTATGAGAAATCCAGTTTGAAGTGTGGGTTGAGAGCTGACAGAGATGCTG 3300  
Db 3241 GTCCTGGTGTATGAGAAATCCAGTTTGAAGTGTGGGTTGAGAGCTGACAGAGATGCTG 3300  
Qy 3301 GACACCGAGAAAACAGAGCAGGCGGAGCCGATCAGCGATCACCGAGTCTCGCAGGTG 3360  
Db 3301 GACACCGAGAAAACAGAGCAGGCGGAGCCGATCAGCGATCACCGAGTCTCGCAGGTG 3360  
Qy 3361 GTGAGCTGGCAGTGAAGGAGCACAAGGCTGAGATTCTCGCTCTGACGAGGCTCTCAAA 3420  
Db 3361 GTGAGCTGGCAGTGAAGGAGCACAAGGCTGAGATTCTCGCTCTGACGAGGCTCTCAAA 3420  
Qy 3421 GAGCAGAAGCTGAAGCGGAGAGCCTCTCTGACAAAGCTCAATGACCTGGAGAAAGACAT 3480  
Db 3421 GAGCAGAAGCTGAAGCGGAGAGCCTCTCTGACAAAGCTCAATGACCTGGAGAAAGACAT 3480  
Qy 3481 GCTATGCTGAAATGAATGCTCCGAGGCTTACAGCAGAGCTGGAGACTGACAGAGCTC 3540  
Db 3481 GCTATGCTGAAATGAATGCTCCGAGGCTTACAGCAGAGCTGGAGACTGACAGAGCTC 3540  
Qy 3541 AAACAGAGGCTTCTGGAGAGCAAGCCAAATTTACAGCAGAGATGACCTGCGAGAAAAT 3600  
Db 3541 AAACAGAGGCTTCTGGAGAGCAAGCCAAATTTACAGCAGAGATGACCTGCGAGAAAAT 3600  
Qy 3601 CACATTTTCGCTGACTCAAGGACTGCAAGAGCTCTAGATCGGCTGATCTACTGAAG 3660  
Db 3601 CACATTTTCGCTGACTCAAGGACTGCAAGAGCTCTAGATCGGCTGATCTACTGAAG 3660  
Qy 3661 ACAGAAAGAGTGAATTTGGAGTATCAGCTGGAACAACTTCAAGTCTCTATTTCTCATGAA 3720  
Db 3661 ACAGAAAGAGTGAATTTGGAGTATCAGCTGGAACAACTTCAAGTCTCTATTTCTCATGAA 3720  
Qy 3721 AAGTGAAAATGGAAGCAGCTATTTCTCAACAAACCAAACTGATTTTCTGCAAGCC 3780  
Db 3721 AAGTGAAAATGGAAGCAGCTATTTCTCAACAAACCAAACTGATTTTCTGCAAGCC 3780  
Qy 3781 AAATGACCAACTGCTTAAAGAAAGAGGTTTATTTAGTTCGACGGAAGAGGACCT 3840  
Db 3781 AAATGACCAACTGCTTAAAGAAAGAGGTTTATTTAGTTCGACGGAAGAGGACCT 3840  
Qy 3841 GCTTTACCCACAGGTTCTCTGCAATCAATGAGCTGAAGTGGCCCTGGAGAGAGG 3900  
Db 3841 GCTTTACCCACAGGTTCTCTGCAATCAATGAGCTGAAGTGGCCCTGGAGAGAGG 3900  
Qy 3901 AAAGCTCGCTGTCAGAGCTAGAGAGCCCTTCAGAGACCCGATCGAGCTCCGCTC 3960  
Db 3901 AAAGCTCGCTGTCAGAGCTAGAGAGCCCTTCAGAGACCCGATCGAGCTCCGCTC 3960  
Qy 3961 GCCCGGAGGAGGCTGCCACCGCAAGCAACGAGCAACCCACACCCATCCAGCCAGCC 4020  
Db 3961 GCCCGGAGGAGGCTGCCACCGCAAGCAACGAGCAACCCACACCCATCCAGCCAGCC 4020



Db 6181 GCAGAGTTCATGCTTCTAGACGTGGTGACCTTAAATAAGCCCTTAAGCGTCGAGGC 6240  
QY 6241 CAGCCACCTCTCTCTACAAAAGAGTACTAGTGCACATGACTGTGAAGAAACAATGTGAA 6300  
Db 6241 CAGCCACCTCTCTCTACAAAAGAGTACTAGTGCACATGACTGTGAAGAAACAATGTGAA 6300  
QY 6301 AACCTCATCTGAATCTGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCC 6360  
Db 6301 AACCTCATCTGAATCTGAAAGCTTCTAATTTCTATAGAAATGACACCTCCCTGGAGCC 6360  
QY 6361 GAGAGACAATCTGTTGTTGATTTTGAAGACAGGCAAGCAACACTGTATTTAGTTCCA 6420  
Db 6361 GAGAGACAATCTGTTGTTGATTTTGAAGACAGGCAAGCAACACTGTATTTAGTTCCA 6420  
QY 6421 TAGCCAGGCTCAACAGGACAGTGGCTGGGCTTTAAAAACACACAGATGACTGAAATG 6480  
Db 6421 TAGCCAGGCTCAACAGGACAGTGGCTGGGCTTTAAAAACACACAGATGACTGAAATG 6480  
QY 6481 ATGTGTGGCTCAGTCCCTCTTCCAGAAATTTTACTGGCAAGGAGTTAGCATTCATT 6540  
Db 6481 ATGTGTGGCTCAGTCCCTCTTCCAGAAATTTTACTGGCAAGGAGTTAGCATTCATT 6540  
QY 6541 TTGGCTTAAGAAAAATCGAAGATGATGTTTGA 6574  
Db 6541 TTGGCTTAAGAAAAATCGAAGATGATGTTTGA 6574

## RESULT 2

AAD38864  
ID AAD38864 standard; cDNA; 6298 BP.

XX AAD38864;

XX 23-SEP-2002 (first entry)

XX Human kinase (PKIN)-21 cDNA.

XX Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;  
KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;  
KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;  
KW development; hepatitis; cardiovascular; hypertension; drug screening;  
KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;  
KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;  
KW hypercholesterolemia; obesity; gene therapy; cytostatic; anti-HIV;  
KW neuroprotective; hepatotropic; hypotensive; cardiac; nephrotropic;  
KW hyperlipidaemia; enzyme; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH 55..6219  
CDS /\*tag= a

FT /product= "Human kinase (PKIN)-21"

XX WO200233099-A2.

XX 25-APR-2002.

XX 20-OCT-2001; 2001WO-US047728.

XX 20-OCT-2000; 2000US-0242410P.

XX 27-OCT-2000; 2000US-0244068P.

XX 03-NOV-2000; 2000US-0245708P.

XX 09-NOV-2000; 2000US-0247672P.

XX 16-NOV-2000; 2000US-0249565P.

XX 22-NOV-2000; 2000US-0252730P.

XX 01-DEC-2000; 2000US-0250807P.

XX (INCY-) INCYTE GENOMICS INC.

XX Gururajan R, Baughn MR, Walia NK, Elliott VS, Xu Y, Arvizu C;

PI Yao MG, Ramkumar J, Ding L, Tang YT, Hafalia AJA, Nguyen DB;

PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;

PI Recipon SA, Lu DAM, Borowsky ML, Thornton M, Swarnaker A;  
PI Thangavelu K, Khan FA, Ison CH;  
XX WPI; 2002-454603/48.  
DR P-PSDB; AAE24150.

XX New human kinase polypeptide, for diagnosing, preventing and treating  
PT cancer, immune system disorders, growth and development disorders,  
PT cardiovascular disorders and lipid disorders.

XX Claim 5; Page 207-209; 210pp; English.

XX The invention relates human kinases (PKIN) and their corresponding  
CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,  
CC treating and preventing cancer, an immune system disorder (e.g., acquired  
CC immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,  
CC atherosclerosis, multiple sclerosis, psoriasis), disorders affecting  
CC growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),  
CC cardiovascular disorder (e.g., hypertension, myocardial infarction,  
CC Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,  
CC Gaucher's disease, Niemann-Pick's disease, hypercholesterolemia,  
CC hyperlipidaemia, obesity), and for assessing the effects of exogenous  
CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a  
CC condition or a disease associated with the expression of PKIN in a  
CC biological sample. A composition comprising PKIN or an agonist or  
CC antagonist of PKIN is useful for treating a disease or condition  
CC associated with decreased or increased expression of functional PKIN.  
CC PKIN is useful in a number of drug screening techniques and to analyse  
CC the proteome of a tissue or cell type. PKIN DNA is useful for creating  
CC knockin humanised animals or transgenic animals to model human diseases,  
CC and in somatic or germline gene therapy. The present sequence is human  
CC PKIN cDNA

XX Sequence 6298 BP; 1772 A; 1585 C; 1720 G; 1221 T; 0 U; 0 Other;

XX Query Match 86.5%; Score 5684.4; DB 6; Length 6298;

XX Best Local Similarity 95.5%; Pred. No. 0;

XX Matches 5979; Conservative 0; Mismatches 36; Indels 243; Gaps 3;

QY 1 AGAGCCGCCAGTGGGAGATGTTGAAGTTCAATATGGAGCGGGAATCTTTGGATGCT 60

Db 37 AGAGCCGCCAGTGGGAGATGTTGAAGTTCAATATGGAGCGGGAATCTTTGGATGCT 96

QY 61 GGTGCTCTGAACCCATTGTCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAA 120

Db 97 GGTGCTCTGAACCCATTGTCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAA 156

QY 121 CCACCTTTATGACTCAACAGAGATGTCCTCTTTCCGAGAAGGATATTAGATGCC 180

Db 157 CCACCTTTATGACTCAACAGAGATGTCCTCTTTCCGAGAAGGATATTAGATGCC 216

QY 181 CTCTTTGTTCTCTTTGAAGAATGACGTACGCTGCTCTGATGAAGATTAAGCAGTGAGC 240

Db 217 CTCTTTGTTCTCTTTGAAGAATGACGTACGCTGCTCTGATGAAGATTAAGCAGTGAGC 276

QY 241 AACTTTGTCGGAAGTATTCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCA 300

Db 277 AACTTTGTCGGAAGTATTCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCA 336

QY 301 AAGGACTTCGAAGTCAGAAGTCTTTGAGTTGTTGTCACCTTGTGTAAGTCAGGTGTA 360

Db 337 AAGGACTTCGAAGTCAGAAGTCTTTGAGTTGTTGTCACCTTGTGTAAGTCAGGTGTA 396

QY 361 AGAGAGAAAGCAACCGGGGACATCTATGCTATGAAGTGTATGAAGAAGAGGCTTTATTG 420

Db 397 AGAGAGAAAGCAACCGGGGACATCTATGCTATGAAGTGTATGAAGAAGAGGCTTTATTG 456

QY 421 GCCCAGGACAGGTTTTCATTTTTTGGAGAGCGGGAACATATTATCTCAAGCAACAGC 480

Db 457 GCCCAGGACAGGTTTTCATTTTTTGGAGAGCGGGAACATATTATCTCAAGCAACAGC 516

QY 481 CCGTGGATCCCCCAATTACAGTATGCTTTTCAGGACAAAATCACCTTTATCTCTGTGAG 540





Db 4852 GAAGCTGATGCTAAACTGCTTGGAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTTA 4911  
QY 4723 GACATGAATGACGCTGCCCTTCAGTGACACAGTGGTGTGGTGGCCACCGAGGAGG 4782  
Db 4912 GACATGAATGACGCTGCCCTTCAGTGACACAGTGGTGTGGTGGCCACCGAGGAGG 4971  
QY 4783 CTCTACGCCCTGAATGCTTTGAAAACTCCCTAACCCATGTCCTCCAGGAATTTGAGCAGTC 4842  
Db 4972 CTCTACGCCCTGAATGCTTTGAAAACTCCCTAACCCATGTCCTCCAGGAATTTGAGCAGTC 5031  
QY 4843 TTCCAAATTTATATATCAAGGACCTGGAGAACTACTCATGATAGCAGGAGAGCGG 4902  
Db 5032 TTCCAAATTTATATATCAAGGACCTGGAGAACTACTCATGATAGCAGGAGAGCGG 5091  
QY 4903 GCATGCTGCTTTGGACGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCT 4962  
Db 5092 GCATGCTGCTTTGGACGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCT 5151  
QY 4963 GCCAGCCGACATCTCACCAACATTTTGAAGCTGTCAAGGCTGGCCACTTTTGG 5022  
Db 5152 GCCAGCCGACATCTCACCAACATTTTGAAGCTGTCAAGGCTGGCCACTTTTGG 5211  
QY 5023 GCAGGCAAGATTTGAGAACGGCTCTGCATCTGTGCAGCCATGCCAGCAAGTCTGCTATT 5082  
Db 5212 GCAGGCAAGATTTGAGAACGGCTCTGCATCTGTGCAGCCATGCCAGCAAGTCTGCTATT 5271  
QY 5083 CTCGCTACCAAGAAACCTCAGCAATATCTGCATCGGAAAGAGATAGAGACCTCAGAG 5142  
Db 5272 CTCGCTACCAAGAAACCTCAGCAATATCTGCATCGGAAAGAGATAGAGACCTCAGAG 5331  
QY 5143 CCTGAGCTGTATCCACTTCACTTACCAATACAGTATCTCATTTGGAACCAATAATCTAC 5202  
Db 5332 CCTGAGCTGTATCCACTTCACTTACCAATACAGTATCTCATTTGGAACCAATAATCTAC 5391  
QY 5203 GAAATGCACATGAAGCAGTACAGCTCGAGGAATTTCTGTAAGAAATGACCAATTCCTTG 5262  
Db 5392 GAAATGCACATGAAGCAGTACAGCTCGAGGAATTTCTGTAAGAAATGACCAATTCCTTG 5451  
QY 5263 GCACCTGCTGTTTGGCCGCTTTCACACAGCTTCCCTGCTCAATCTGCAGGTGAAC 5322  
Db 5452 GCACCTGCTGTTTGGCCGCTTTCACACAGCTTCCCTGCTCAATCTGCAGGTGAAC 5511  
QY 5323 AGCGAGGCGAGGAGGAGTACTTCTGCTGTTTCCACGAATTTGAGTGTTCGTGGAT 5382  
Db 5512 AGCGAGGCGAGGAGGAGTACTTCTGCTGTTTCCACGAATTTGAGTGTTCGTGGAT 5571  
QY 5383 TCTTAGGAAAGAGTACGCGCAGACGATCTCAAGTGAGTGCCTTACCTTTGGCCTTT 5442  
Db 5572 TCTTAGGAAAGAGTACGCGCAGACGATCTCAAGTGAGTGCCTTACCTTTGGCCTTT 5631  
QY 5443 GCCTACAGAAACCTTATCTGTTTGTGACCACTTCAACTCTCAAGTGAATTTGAGATC 5502  
Db 5632 GCCTACAGAAACCTTATCTGTTTGTGACCACTTCAACTCTCAAGTGAATTTGAGATC 5691  
QY 5503 CAGGCAAGCTCTCAGCAGGACCCCTGCGCAGCGTACTCGACATCCCGAACCCGCGC 5562  
Db 5692 CAGGCAAGCTCTCAGCAGGACCCCTGCGCAGCGTACTCGACATCCCGAACCCGCGC 5751  
QY 5563 TAGCTGGCCCTGCCATTTCTCAGAGGATTTACTTGGCTGCTCATACAGGATAA 5622  
Db 5752 TAGCTGGCCCTGCCATTTCTCAGAGGATTTACTTGGCTGCTCATACAGGATAA 5811  
QY 5623 TTAAGGCTCATTTGCTGCAAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACCCCGG 5682  
Db 5812 TTAAGGCTCATTTGCTGCAAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACCCCGG 5871  
QY 5683 GGCCGCTCAGCTCCGCGACAGCCCAACAGCGAGGCGCCACCCAGTACACAGCAC 5742  
Db 5872 GGCCGCTCAGCTCCGCGACAGCCCAACAGCGAGGCGCCACCCAGTACACAGCAC 5931  
QY 5743 ATCACAAGCGCTGCTCAGCCAGGCGCCGCGAGGCGCCAGGCGCCAGCCACCCGAGAG 5802  
Db 5932 ATCACAAGCGCTGCTCAGCCAGGCGCCGCGAGGCGCCAGGCGCCAGCCACCCGAGAG 5991

QY 5803 CCAAGCACACCCACCGCTACCGGAGGGCGGACCGAGCTGCGCAGGCAAGTCTCCT 5862  
Db 5992 CCAAGCACACCCACCGCTACCGGAGGGCGGACCGAGCTGCGCAGGCAAGTCTCCT 6051  
QY 5863 GGCCGCGCCCTGGAGCGAGAGAGTCCCGCGGATGCTCAGCAGCGGAGAGCGG 5922  
Db 6052 GGCCGCGCCCTGGAGCGAGAGAGTCCCGCGGATGCTCAGCAGCGGAGAGCGG 6111  
QY 5923 TCCCGCGGAGGCTGTTTGAAGACAGCAGCAGGCGGCGGCTGCTGCGGAGCGCTGAG 5982  
Db 6112 TCCCGCGGAGGCTGTTTGAAGACAGCAGCAGGCGGCGGCTGCTGCGGAGCGCTGAG 6171  
QY 5983 ACCCGCTGTCCTCCAGGTGAACAGGAGGCGGAGTGCCTCTCAAGTTTTCACGGTT 6042  
Db 6172 ACCCGCTGTCCTCCAGGTGAACAGGAGTGCCTGCGGACCGAGTCTCAGTATTAATCTCAGCCAGA 6231  
QY 6043 AACACTGTCACCTATTAT 6060  
Db 6232 AAAACCAACTCCTCATCT 6249

## RESULT 3

AAL55215

ID AAL55215 standard; DNA; 8603 BP.

AC AAL55215;

XX  
DT 01-MAY-2003 (first entry)XX  
DE Human CR1K related DNA sequence, SEQ ID No 4.

XX  
KW Anorectic; hypotensive; cardiac; antilipemic; cerebroprotective;  
KW angiot; osteopathic; antiarthritic; cytostatic; antidepressant;  
KW immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;  
KW neuroprotective; antiinflammatory; antidiabetic; analgesic;  
KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;  
KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;  
KW central nervous system disorder; chronic obstructive pulmonary disease;  
KW diabetes; pain; ds.

OS Homo sapiens.

XX WO2003004523-A1.

XX  
PD 16-JAN-2003.XX  
PF 28-JUN-2002; 2002WO-EP007156.XX  
PR 02-JUL-2001; 2001US-0301841P.XX  
PR 11-DEC-2001; 2001US-0338651P.XX  
PR 25-APR-2002; 2002US-0375014P.XX  
PA (FARB ) BAYER AG.XX  
PI Zhu Z;XX  
PI WPI; 2003-221576/21.

XX  
PT New human citron rho/rac-interacting kinase (CR1K) polypeptide and  
PT polynucleotide, useful in preventing, ameliorating or treating diseases  
PT associated with human CR1K dysfunction, e.g. obesity, diabetes or  
PT Alzheimer's disease.

XX  
PS Disclosure; Fig 4; 237pp; English.

XX  
CC The invention relates to an isolated polynucleotide encoding a human  
CC citron rho/rac-interacting kinase polypeptide. The isolated  
CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the  
CC specification. The human citron rho/rac-interacting kinase (CR1K)  
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or  
CC treating diseases associated with human CR1K dysfunction, such as obesity  
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery

CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of  
CC cancer including endometrial, breast, prostate and colon cancer),  
CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood  
CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's  
CC disease), chronic obstructive pulmonary disease, or diabetes. These can  
CC also be used to treat pain associated with the disorders. The human CR1K  
CC polypeptide is also useful in diagnostic assays or in genetic testing.  
CC The expression vector or the reagent is useful in preparing a medicament  
CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a  
CC central nervous system disorder, or chronic obstructive pulmonary  
CC disease. The fusion protein is useful for generating antibodies against a  
CC CR1K polypeptide and for use in various assay systems. The methods are  
CC useful in producing and detecting the polynucleotide and polypeptide and  
CC in screening for agents that modulate the activity of the human CR1K  
CC polypeptide. This polynucleotide sequence represents a DNA sequence  
CC relating to the human CR1K protein of the invention  
XX  
SQ Sequence 8603 BP; 2305 A; 2206 C; 2215 G; 1877 T; 0 U; 0 Other;

Query Match 86.2%; Score 5668; DB 8; Length 8603;  
Best Local Similarity 95.5%; Pred No. 0;  
Matches 5962; Conservative 0; Mismatches 35; Indels 243; Gaps 3;

QY 19 ATGTTGAAGTTCAATATAGAGCGGGGAATCCTTTGGATGCTGGTGTGCTGTAACCCCAATT 78  
DB 1 ATGTTGAAGTTCAATATAGAGCGGGGAATCCTTTGGATGCTGGTGTGCTGTAACCCCAATT 60  
QY 79 GCACGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAACCAACCTTTATGACTCAA 138  
DB 61 GCACGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAACCAACCTTTATGACTCAA 120  
QY 139 CAGCAGATGTCCTCTTTCCGAGAAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198  
DB 121 CAGCAGATGTCCTCTTTCCGAGAAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
QY 199 GAATGAGTACGCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 258  
DB 181 GAATGAGTACGCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
QY 259 TCCGACACCATAGCTGAGTTACAGAGCTCCAGGCTTCGGCAAGGACTTCGAAGTCAGA 318  
DB 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGGCTTCGGCAAGGACTTCGAAGTCAGA 300  
QY 319 AGTCTTGATGTTGGTGCATTTGCTGAACTGATGAGAGAGGCTTTATTGGCCCGAGCAGGTTCA 438  
DB 301 AGTCTTGATGTTGGTGCATTTGCTGAACTGATGAGAGAGGCTTTATTGGCCCGAGCAGGTTCA 420  
QY 379 GACATCTATGCTATGAAAGTATGAGAGAGGCTTTATTGGCCCGAGCAGGTTCA 498  
DB 361 GACATCTATGCTATGAAAGTATGAGAGAGGCTTTATTGGCCCGAGCAGGTTCA 480  
QY 439 TTTTGTGAGGAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 498  
DB 421 TTTTGTGAGGAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
QY 499 CAGTATGCCCTTTCCAGACAAAATCACCTTTATCTGATGAGGAATATCAGCCTGGAGGG 558  
DB 481 CAGTATGCCCTTTCCAGACAAAATCACCTTTATCTGATGAGGAATATCAGCCTGGAGGG 540  
QY 559 GACTTCTGCTGCACTTTTGAATAGATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 618  
DB 541 GACTTCTGCTGCACTTTTGAATAGATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600  
QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGTTTCATCTGATGGGATACGTGATCGA 578  
DB 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGTTTCATCTGATGGGATACGTGATCGA 660  
QY 679 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTT 738  
DB 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGAATTT 720  
QY 739 GGATCTGCCGCGAAAATGAATTCAAACAAAGATGGTGAATGCCAAATCCCGAATGGGACC 798  
DB 1801 CGGAAAGCGAGAAATGTCAGCATAAAATCTGTGAAAGGCTAAGGATCAAGGAAAGCCCTGAA 1860

DB 721 GGATCTGCCGCGAAAATGAATTCAAACAAAGATGGTGAATGCCAAATCCCGATTGGGACC 780  
QY 799 CCAGATTACATGGCTCTCTGAAGTCTGACTGTGATGAACCGGGATGAAAGAGCACCTAC 858  
DB 781 CCAGATTACATGGCTCTCTGAAGTCTGACTGTGATGAACCGGGATGAAAGAGCACCTAC 840  
QY 859 GGCTGAGACTGTGACTGCTGCTCAGTGGGCTGATTGCTCTATGAGATGATTTATGGGAGA 918  
DB 841 GGCTGAGACTGTGACTGCTGCTCAGTGGGCTGATTGCTCTATGAGATGATTTATGGGAGA 900  
QY 919 TCCCTCTTCGAGAGGGAACTCTGCGAGAACCTTCAATAACATTTATGAATTTCCAGCGG 978  
DB 901 TCCCTCTTCGAGAGGGAACTCTGCGAGAACCTTCAATAACATTTATGAATTTCCAGCGG 960  
QY 979 TTTTGAATTTCCAGATGACCCCAAGTGCAGAGTGACTTTCTTGTATCTGATCTGATTCAGAGC 1038  
DB 961 TTTTGAATTTCCAGATGACCCCAAGTGCAGAGTGACTTTCTTGTATCTGATTCAGAGC 1020  
QY 1039 TTTGTTGTCGCGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTCTGCTGCCATCTCTTCTTC 1098  
DB 1021 TTTGTTGTCGCGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTCTGCTGCCATCTCTTCTTC 1080  
QY 1099 TCTRAAATTCAGTGGAAACATTCGTAACCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1158  
DB 1081 TCTRAAATTCAGTGGAAACATTCGTAACCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1140  
QY 1159 TCTGACGATGACCTCCAAATTTTGTGAACACAGAGAGAAATTCGTGGGTTTTCATCTCTCT 1218  
DB 1141 TCTGACGATGACCTCCAAATTTTGTGAACACAGAGAGAAATTCGTGGGTTTTCATCTCTCT 1200  
QY 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGGGGTTTTTCG 1278  
DB 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGGGGTTTTTCG 1260  
QY 1279 TACAGCAAGCAGCTGGGGATTCTTTGGTAGATCTGAGTCTGTGTGTCGGGTCTGGAGTCC 1338  
DB 1261 TACAGCAAGCAGCTGGGGATTCTTTGGTAGATCTGAGTCTGTGTGTCGGGTCTGGAGTCC 1320  
QY 1339 CTTCCCAAGACTAGCTCCATCGAAAGAAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1398  
DB 1321 CTTCCCAAGACTAGCTCCATCGAAAGAAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1380  
QY 1399 TCTCAGCAAGTGTCTCAGAGTGGAGCAAGAAATGACCCGTTTACATCCGAGAGTGTCA 1458  
DB 1381 TCTCAGCAAGTGTCTCAGAGTGGAGCAAGAAATGACCCGTTTACATCCGAGAGTGTCA 1440  
QY 1459 GAGTGGAGGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518  
DB 1441 GAGTGGAGGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
QY 1519 TCCCTCTGAGAGGAGCTTGTCTTACCTATCAGAAATGAGAGTCTTAAAGCGAAGT 1578  
DB 1501 TCCCTCTGAGAGGAGCTTGTCTTACCTATCAGAAATGAGAGTCTTAAAGCGAAGT 1560  
QY 1579 TTGAGCAAGCAGGATGGAGGTGTCTCCAGAGGATGCAAAAGCACTGCGAGCTTCTCCAT 1638  
DB 1561 TTGAGCAAGCAGGATGGAGGTGTCTCCAGAGGATGCAAAAGCACTGCGAGCTTCTCCAT 1620  
QY 1639 GATATCAGAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1698  
DB 1621 GATATCAGAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1680  
QY 1699 GTGGAAGAAATGAGGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1758  
DB 1681 GTGGAAGAAATGAGGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740  
QY 1759 CGGAGTGTCTTACGAATCTGAGCTGAGAGGTTCTCGGTTGCTGCTGAAGAAATTCAG 1818  
DB 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGGTTCTCGGTTGCTGCTGAAGAAATTCAG 1800  
QY 1819 CGGAAAGCGAGAGATGTCAGCATAAAATCTTGAAGCTAAGGATCAAGGAGACCTGAA 1878  
DB 1801 CGGAAAGCGAGAGATGTCAGCATAAAATCTTGAAGGCTAAGGATCAAGGAGACCTGAA 1860

1879 GTGGAGATATATCGGAACTTGGAGAGATCAATGCTTGAGCAGCAGCTCAAAATTCAGGAG 1938  
1881 GTGGAGATATATCGGAACTTGGAGAGATCAATGCTTGAGCAGCAGCTCAAAATTCAGGAG 1920  
1939 CTCGACAGAACTTGGAGAGGCTG----- 1963  
1921 CTCGAGAGAACTTGGAGAGGCTGTTAAAGCCAGCAGCGAGGCCACCGAGCTGCTGCGAG 1980  
1964 -----CAAGAGCCGAGCGAGGAGGAGCTGGAGAGCTGGAGAGCTGCGAGAACCGAGAG 2010  
1981 AATATCCGCGCAGGCAAGAGGAGCGAGGAGGAGCTGGAGAGCTGCGAGAACCGAGAG 2040  
2011 GATTCCTTCAAGGCTACAGAAAGAGCTGGTGGAGAGCTGAGGAAACCGCCCATTCCTG 2070  
2041 GATTCCTTCAAGGCTACAGAAAGAGCTGGTGGAGAGCTGAGGAAACCGCCCATTCCTG 2100  
2071 GAGAACCAAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGACTGGAAGATGAC 2130  
2101 GAGAACCAAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGACTGGAAGATGAC 2160  
2131 ATCCAGACAAATCCCAAGATCCAGCAGATCGCTGATATAAATTCCTGGAGCTCGAGAG 2190  
2161 ATCCAGACAAATCCCAAGATCCAGCAGATCGCTGATATAAATTCCTGGAGCTCGAGAG 2220  
2191 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAAAGTGCACCTGAAACAGAAAGAG 2250  
2221 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAAAGTGCACCTGAAACAGAAAGAG 2280  
2251 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATTAAGAAAGACCTGGCTGAC 2310  
2281 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATTAAGAAAGACCTGGCTGAC 2340  
2311 AAGGAGACCTGAGAAACATGATGAGAGACAGAGGAGGAGCCCATGAGAGGSCAAA 2370  
2341 AAGGAGACCTGAGAAACATGATGAGAGACAGAGGAGGAGCCCATGAGAGGSCAAA 2400  
2371 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTPATGGATTCCAAGATTCAGATCCCTGGAA 2430  
2401 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTPATGGATTCCAAGATTCAGATCCCTGGAA 2460  
2431 CAGAGGATGTGAACTGTCTGAAGCCCAATTAAGTGTGAGCAATAGCAGTCTTTTACC 2490  
2461 CAGAGGATGTGAACTGTCTGAAGCCCAATTAAGTGTGAGCAATAGCAGTCTTTTACC 2520  
2491 CAAAGGAACATGAGGCCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550  
2521 CAAAGGAACATGAGGCCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580  
2551 CTGAGACACAGCTGGGAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAG 2610  
2581 CTGAGACACAGCTGGGAGTTGGAGGCCAGAACCGGAAACTGGAGGAGCAGCTGGAG 2640  
2611 AAGATCAGCCACCAAGACACAGTGAAGAACTGGCTGTGAACTGGAGACAAGATTG 2670  
2641 AAGATCAGCCACCAAGACACAGTGAAGAACTGGCTGTGAACTGGAGACAAGATTG 2700  
2671 CGGAGGTGAGTCTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2730  
2701 CGGAGGTGAGTCTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2760  
2731 CTACAGCTCTCCCTCGAGAGCGCAGTCAAGTTCACAGCCCTGAGGCTCAGCGGCG 2790  
2761 CTACAGCTCTCCCTCGAGAGCGCAGTCAAGTTCACAGCCCTGAGGCTCAGCGGCG 2820  
2791 GCCTTGGAGGCGAGTTTCGCCAGGCGAAGACAGAGCTGGAGAGACCAAGCAGAGCT 2850  
2821 GCCTTGGAGGCGAGTTTCGCCAGGCGAAGACAGAGCTGGAGAGACCAAGCAGAGCT 2880  
2851 GAAGAGGAGATCCAGGCACTCAGGCAACATAGAGATGAATCCAGGCGAAATTTGATGCT 2910  
2881 GAAGAGGAGATCCAGGCACTCAGGCAACATAGAGATGAATCCAGGCGAAATTTGATGCT 2940

2911 CTTGTAACAGCTGTACTGTAAATCACAGACTTGAGGAGCAGCTAAACCAGCTGACCGAG 2970  
2941 CTTGTAACAGCTGTACTGTAAATCACAGACTTGAGGAGCAGCTAAACCAGCTGACCGAG 3000  
2971 GACAAACGCTGAACCTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3030  
3001 GACAAACGCTGAACCTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060  
3031 GCGCCAAACGAGAGATTGTAACTCGGAAGTGAAGTGAACCACTCTCCGCGGAGATC 3090  
3061 GCGCCAAACGAGAGATTGTAACTCGGAAGTGAAGTGAACCACTCTCCGCGGAGATC 3120  
3091 ACGGAAACGAGAGATGACAGCTTACCAGCAGAGCAAGCAACGATGGAGCTCTGAAGACCAAG 3150  
3121 ACGGAAACGAGAGATGACAGCTTACCAGCAGAGCAAGCAACGATGGAGCTCTGAAGACCAAG 3180  
3151 TGCACCAATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA 3210  
3181 TGCACCAATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTAAACGATGAGCTGCTAGAA 3240  
3211 AAGAGCGGAGTGGAGGCTGGAGGAGCGTCTGGTGTGATGAGAAATCCAGATTTGAG 3270  
3241 AAGAGCGGAGTGGAGGCTGGAGGAGCGTCTGGTGTGATGAGAAATCCAGATTTGAG 3300  
3271 TGTCCGCTTCCAGAGCTGCGAGAAATCTGACACCCAGAGAAACAGAGCGGCGAGAGCC 3330  
3301 TGTCCGCTTCCAGAGCTGCGAGAAATCTGACACCCAGAGAAACAGAGCGGCGAGAGCC 3360  
3331 GATCAGCGGATCACCGAGTCTCCGAGCTGGTGGAGCTGGAGTGGAGGAGCAGAGCT 3390  
3361 GATCAGCGGATCACCGAGTCTCCGAGCTGGTGGAGCTGGAGTGGAGGAGCAGAGCT 3420  
3391 GAGATTCTCGCTCTGCGAGCAGGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3450  
3421 GAGATTCTCGCTCTGCGAGCAGGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3480  
3451 GACAACTCAATGACCTGGAGAGAGCATGCTATGTTGAAATGAATGCCGAGCTTA 3510  
3481 GACAACTCAATGACCTGGAGAGAGCATGCTATGTTGAAATGAATGCCGAGCTTA 3540  
3511 CAGCAGAGCTGGAGACTGAACAGAGCTCAAAACAGAGGCTTCTGAAGAGCAGAGCAAA 3570  
3541 CAGCAGAGCTGGAGACTGAACAGAGCTCAAAACAGAGGCTTCTGAAGAGCAGAGCAAA 3600  
3571 TTACAGCAGCAGATGAGACCTGCGAGAAATCAATTTTCGCTGTGACTCAAGGACTGCAA 3630  
3601 TTACAGCAGCAGATGAGACCTGCGAGAAATCAATTTTCGCTGTGACTCAAGGACTGCAA 3660  
3631 GAAGCTCTAGATCGGCTGATCTACTGAACAGAGAGAGTGAAGTGGAGTATCAGCTG 3690  
3661 GAAGCTCTAGATCGGCTGATCTACTGAACAGAGAGAGTGAAGTGGAGTATCAGCTG 3720  
3691 GAAACCAATTCAGGTTCTTATTTCTCATGAAAGGTGAAATGGAAGCAGCTATTTCTCAA 3750  
3721 GAAACCAATTCAGGTTCTTATTTCTCATGAAAGGTGAAATGGAAGCAGCTATTTCTCAA 3780  
3751 CAAACCAATTCAGTGTGATTTCTGCAAGCCAAATGGAACCACTGCTGAAAGAGAAAG 3810  
3781 CAAACCAATTCAGTGTGATTTCTGCAAGCCAAATGGAACCACTGCTGAAAGAGAAAG -- 3838  
3811 GGTTTATTTAGTCGAGGAAAGAGGACCTGCTTTACCCACAGAGTTCCTCTGAGTAC 3870  
3839 -----AGTTCTCTGAGTAC 3855  
3871 AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAGTCTGCTGAGAGCTAGAGGAGCC 3930  
3856 AATGAGCTGAAGCTGGCCCTGGAGAGAGAGAGTCTGCTGAGAGCTAGAGGAGCC 3915  
3931 CTTCAGAGAGCCGAGCTCAGAGCTCCGCTCCGCGGAGGAGAGCTGCCACCCGCAAGCA 3990  
3916 CTTCAGAGAGCCGAGCTCAGAGCTCCGCTCCGCGGAGGAGAGCTGCCACCCGCAAGCA 3975  
3991 ACGGACCAACCAACCCATCCAGCCAGCAACCCGCGGAGGAGAGATCGCCATGTCGCC 4050

Db 3976 ACAGACCCACCCATCCAGCCAGCCACCGGAGGACAGATGCCATGTCCGCC 4035  
Qy 4051 ATCTGTGCGGTGCCAGAGACACAGCCCACTGAGCCCTGCTGGCCCCCACCAC 4110  
Db 4036 ATCTGTGCGGTGCCAGAGACACAGCCCACTGAGCCCTGCTGGCCCCCACCAC 4095  
Qy 4111 CCAGAGAAAGGAGTCTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAGGACGATGCAC 4170  
Db 4096 CCAGAGAAAGGAGTCTCAACTCCAGAGGAATTTAGTCGGCGCTTTAAGGACGATGCAC 4155  
Qy 4171 CACAATATTCCTCACCGGATTCACACCTAGGACTGAACTGCGAGCCACAAAGTGTCTGTG 4230  
Db 4156 CACAATATTCCTCACCGGATTCACACCTAGGACTGAACTGCGAGCCACAAAGTGTCTGTG 4215  
Qy 4231 TGTCTGATACCGTGCATTTGGACGCCAGGCATCCAAATGTCCTGAAATGTCAGGTGATG 4290  
Db 4216 TGTCTGATACCGTGCATTTGGACGCCAGGCATCCAAATGTCCTGAAATGTCAGGTGATG 4275  
Qy 4291 TGTCTACCCCAAGTGTCCACAGTGTTCGACGCACCTGCGGCTTCCCTGCTGAAATATGCC 4350  
Db 4276 TGTCTACCCCAAGTGTCCACAGTGTTCGACGCACCTGCGGCTTCCCTGCTGAAATATGCC 4335  
Qy 4351 ACACACTTACCGAGGCTTTGCGGTGACAAATGAACTCCCAAGTCTCCAGACCAAG 4410  
Db 4336 ACACACTTACCGAGGCTTTGCGGTGACAAATGAACTCCCAAGTCTCCAGACCAAG 4395  
Qy 4411 GAGCCACAGCAGCTTCACCTGGAAGGTTGATGAGGTGCCAGGAATAACAGCA 4470  
Db 4396 GAGCCACAGCAGCTTCACCTGGAAGGTTGATGAGGTGCCAGGAATAACAGCA 4455  
Qy 4471 GGACAGCAAGGTGGGACAGGAAGTACATTTCTCTGAGGGATCAAAAGTCTCTATTAT 4530  
Db 4456 GGACAGCAAGGTGGGACAGGAAGTACATTTCTCTGAGGGATCAAAAGTCTCTATTAT 4515  
Qy 4531 GACATGAGCCAGACAGAGCTGACAGAGCGGTGGAGAAATTGAGCTGTGCTTCCC 4590  
Db 4516 GACATGAGCCAGACAGAGCTGACAGAGCGGTGGAGAAATTGAGCTGTGCTTCCC 4575  
Qy 4591 GACGGGGATGATCTATTATCATGTCGCTGCTCCGAACTCGCAATACAGCCAAA 4650  
Db 4576 GACGGGGATGATCTATTATCATGTCGCTGCTCCGAACTCGCAATACAGCCAAA 4635  
Qy 4651 GCA----- 4653  
Db 4636 GCAGATGTCCTACATACATACTGAAGTGAATCTCACCGGCACACACCTGCTGGCCCGG 4695  
Qy 4654 ----- 4653  
Db 4696 AGAACCTCTACTTGTAGCTCCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCCCTTA 4755  
Qy 4654 -----GAAAAAGCAGAAAGCTGATGCTAAACTG 4680  
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAAAAGCAGAAAGCTGATGCTAAAAC 4815  
Qy 4681 CTTGAAACTCCCTGCTGAACTGAAAGTGAATGACCGTCTAGACATGAATGACGCTG 4740  
Db 4816 CTTGAAACTCCCTGCTGAACTGAAAGTGAATGACCGTCTAGACATGAATGACGCTG 4875  
Qy 4741 CCGTTCACTGACAGGTGCTGTTGTTGGCCACCGGAGGAGGCTCTACGCCCTGAATGTC 4800  
Db 4876 CCGTTCACTGACAGGTGCTGTTGTTGGCCACCGGAGGAGGCTCTACGCCCTGAATGTC 4935  
Qy 4801 TTGAAAACTCCCTAACCCATGTCACGAAATTGAGCAGTCTTCCAAATTTATATTATC 4860  
Db 4936 TTGAAAACTCCCTAACCCATGTCACGAAATTGAGCAGTCTTCCAAATTTATATTATC 4995  
Qy 4861 AAGGACCTGGAAGTACTCATATAGCAGAGAGAGCGGCACTGTGCTTGTGGAC 4920  
Db 4996 AAGGACCTGGAAGTACTCATATAGCAGAGAGAGCGGCACTGTGCTTGTGGAC 5055  
Qy 4921 GTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCCAGCCCGACATCTCA 4980

Db 5056 GTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACCTGCTGCCAGCCCGACATCTCA 5115  
Qy 4981 CCACAAATTTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCAGCAAGATTGAGAAC 5040  
Db 5116 CCACAAATTTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCAGCAAGATTGAGAAC 5175  
Qy 5041 GGGCTCTGATCTGTGAGCCATGCCAGCAAAAGTGTCTATTCTCGGCTACAAACGAAAAAC 5100  
Db 5176 GGGCTCTGATCTGTGAGCCATGCCAGCAAAAGTGTCTATTCTCGCTACACGAAAAAC 5235  
Qy 5101 CTCAGCAAAATCTGCAATTCGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5160  
Db 5236 CTCAGCAAAATCTGCAATTCGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5295  
Qy 5161 TTCACCAATTTACAGTATCTCTATTGGAACCAATAAATTTTACGAAATCGACATGAAGCAG 5220  
Db 5296 TTCACCAATTTACAGTATCTCTATTGGAACCAATAAATTTTACGAAATCGACATGAAGCAG 5355  
Qy 5221 TACACGCTCAGGAAATTCCTGGAATAAGATGACCAATTCCTTGGCACCTGCTGTGTTGCC 5280  
Db 5356 TACACGCTCAGGAAATTCCTGGAATAAGATGACCAATTCCTTGGCACCTGCTGTGTTGCC 5415  
Qy 5281 GCCTCTTCCAAAGCTTCCCTGTCTCAATGTCGAGGTGAAACGCGAGGGCAGCGAGAG 5340  
Db 5416 GCCTCTTCCAAAGCTTCCCTGTCTCAATGTCGAGGTGAAACGCGAGGGCAGCGAGAG 5475  
Qy 5341 GAGTACTTGTCTGTGTTTCCACGAATTTGGAGTGTCTGTTGATTTCTTACGGAAGACGTAGC 5400  
Db 5476 GAGTACTTGTCTGTGTTTCCACGAATTTGGAGTGTCTGTTGATTTCTTACGGAAGACGTAGC 5535  
Qy 5401 CGCACAGACGATCTCAAGTGGAGTGTGCTTACCTTTTGGCCTTTGCCTACAGAGAACCTTAT 5460  
Db 5536 CGCACAGACGATCTCAAGTGGAGTGTGCTTACCTTTTGGCCTTTGCCTACAGAGAACCTTAT 5595  
Qy 5461 CTGTTTGTGACCCACATTCACCTCAGTGAATTTGAGATCCAGGCACGCTCTCTCAGCA 5520  
Db 5596 CTGTTTGTGACCCACATTCACCTCAGTGAATTTGAGATCCAGGCACGCTCTCTCAGCA 5655  
Qy 5521 GGGACCCCTCCCGAGCGTACCTGGACATCCCGAAACCCGCGTACCTGGGCCCTGCAATT 5580  
Db 5656 GGGACCCCTCCCGAGCGTACCTGGACATCCCGAAACCCGCGTACCTGGGCCCTGCAATT 5715  
Qy 5581 TCCTCAGGACGATTTACTTTGGGCTCTCATACAGGATTAATTAAGGTCATTTGCTGC 5640  
Db 5716 TCCTCAGGACGATTTACTTTGGGCTCTCATACAGGATTAATTAAGGTCATTTGCTGC 5775  
Qy 5641 AAGGGAAACCTCGTAGGAGTCCGGCACTGAACACACCGGGGCCCGTCCACCTCCCGC 5700  
Db 5776 AAGGGAAACCTCGTAGGAGTCCGGCACTGAACACACCGGGGCCCGTCCACCTCCCGC 5835  
Qy 5701 AGCAGCCCCAACAGCGAGGCCCAACCCAGCTAACAGAGACATCAACAGCGCGTGGCC 5760  
Db 5836 AGCAGCCCCAACAGCGAGGCCCAACCCAGCTAACAGAGACATCAACAGCGCGTGGCC 5895  
Qy 5761 TCAGCCACAGCGCCCGGAGGCCACCCAGCCACCCGAGAGCAAGACACACCCACCCG 5820  
Db 5896 TCAGCCACAGCGCCCGGAGGCCACCCAGCCACCCGAGAGCAAGACACACCCACCCG 5955  
Qy 5821 TACCGCAGGGGGCGGACCGAGCTCCGAGGACAAAGTCTCTGCGCCGCCCTTGGAGCGA 5880  
Db 5956 TACCGCAGGGGGCGGACCGAGCTCCGAGGACAAAGTCTCTGCGCCGCCCTTGGAGCGA 6015  
Qy 5881 GAGAGTCCCCCGCCGGATGCTCAGCACCGGAGAGAGCGGTCCCCCGGAGGCTGTTT 5940  
Db 6016 GAGAGTCCCCCGCCGGATGCTCAGCACCGGAGAGAGCGGTCCCCCGGAGGCTGTTT 6075  
Qy 5941 GAGACAGCAGCAGGGGCGGCTGCTGCGGGAGCGGTGAGGACCCCGTGTGCCAGGTG 6000  
Db 6076 GAGACAGCAGCAGGGGCGGCTGCTGCGGGAGCGGTGAGGACCCCGTGTGCCAGGTG 6135  
Qy 6001 AACAGGGAGAGGGCAGAGTGCCTCTCAAGTTTTTACCGTTAAACTGTCACTATTAT 6060  
Db 6136 AACAGGTCTGGGACCAAGTCTTCAGTATAAATCTCAGCCAGAAAAAACCACTCTCATCT 6195

## RESULT 4

AAL55217  
ID AAL55217 standard; DNA; 6156 BP.

XX AC AAL55217;

XX DT 01-MAY-2003 (first entry)

XX DE Human CR1K related DNA sequence, SEQ ID NO 8.

XX KW Anorectic; hypotensive; cardiac; antilipemic; cerebroprotective;  
XX KW antitumor; osteopathic; antiarthritic; cytostatic; antidepressant;  
XX KW immunomodulator; antianemic; tranquilizer; antiparkinsonian; nootropic;  
XX KW neuroprotective; antiinflammatory; antidiabetic; analgesic;  
XX KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;  
XX KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;  
XX KW central nervous system disorder; chronic obstructive pulmonary disease;  
XX KW diabetes; pain; ds.

XX OS Homo sapiens.

XX PN WO2003004523-A1.

XX XX 16-JAN-2003.

XX XX 28-JUN-2002; 2002WO-BP007156.

XX PR 02-JUL-2001; 2001US-0301841P.

XX PR 11-DEC-2001; 2001US-038651P.

XX PR 25-APR-2002; 2002US-0375014P.

XX PA (PARB ) BAYER AG.

XX PI Zhu Z;

XX DR WPI; 2003-221576/21.

XX PT New human citron rho/rac-interacting kinase (CR1K) polypeptide and  
XX PT polynucleotide, useful in preventing, ameliorating or treating diseases  
XX PT associated with human CR1K dysfunction, e.g. obesity, diabetes or  
XX PT Alzheimer's disease.

XX PS Disclosure; Page 217-222; 237pp; English.

XX CC The invention relates to an isolated polynucleotide encoding a human  
XX CC citron rho/rac-interacting kinase polypeptide. The isolated  
XX CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the  
XX CC specification. The human citron rho/rac-interacting kinase (CR1K)  
XX CC polypeptide and polynucleotide are useful in preventing, ameliorating, or  
XX CC treating diseases associated with human CR1K dysfunction such as obesity  
XX CC and obesity-associated comorbidities (e.g. hypertension, coronary artery  
XX CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of  
XX CC cancer including endometrial, breast, prostate and colon cancer),  
XX CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood  
XX CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's  
XX CC disease), chronic obstructive pulmonary disease, or diabetes. These can  
XX CC also be used to treat pain associated with the disorders. The human CR1K  
XX CC polypeptide is also useful in diagnostic assays or in genetic testing.  
XX CC The expression vector or the reagent is useful in preparing a medicament  
XX CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a  
XX CC central nervous system disorder, or chronic obstructive pulmonary  
XX CC disease. The fusion protein is useful for generating antibodies against a  
XX CC CR1K polypeptide and for use in various assay systems. The methods are  
XX CC useful in producing and detecting the polynucleotide and polypeptide and  
XX CC in screening for agents that modulate the activity of the human CR1K  
XX CC polypeptide. This polynucleotide sequence represents a DNA sequence  
XX CC relating to the human CR1K protein of the invention

XX SQ Sequence 6156 BP; 1732 A; 1548 C; 1679 G; 1197 T; 0 U; 0 Other;

Query Match

86.2%; Score 5666.6; DB 8; Length 6156;

Best Local Similarity 95.9%; Pred. No. 0;  
Matches S945; Conservative 0; Mismatches 9; Indels 243; Gaps 3;

QY 19 ATGTTGAAGTTCAAAATATGGAGCCGGAATCTTTTGGATGCTGGTGTCTGTAACCCATT 78

Db 1 ATGTTGAAGTTCAAAATATGGAGCCGGAATCTTTTGGATGCTGGTGTCTGTAACCCATT 60

QY 79 GCCAGCCGGCCCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCCCTTTATGACTCAA 138

Db 61 GCCAGCCGGCCCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCCCTTTATGACTCAA 120

QY 139 CAGCAGATGTCCTCTTTCCCGAGAAAGGATATTAGATGCCCTCTTTTGTCTCTTTGAA 198

Db 121 CAGCAGATGTCCTCTTTCCCGAGAAAGGATATTAGATGCCCTCTTTTGTCTCTTTGAA 180

QY 199 GAATGTCAGTCAGCTGCTCTGATGAAGATTAGCAGCTGAGCACTTTGTCCGGAAGTAT 258

Db 181 GAATGTCAGTCAGCTGCTCTGATGAAGATTAGCAGCTGAGCACTTTGTCCGGAAGTAT 240

QY 259 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAAGGACTTCGAAAGTCAGA 318

Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAAGGACTTCGAAAGTCAGA 300

QY 319 AGTCTTGTAGTTGTGTCCTTTGTCGAAGTCAGGTCGTAAGAGAGAAAGCAACCGGG 378

Db 301 AGTCTTGTAGTTGTGTCCTTTGTCGAAGTCAGGTCGTAAGAGAGAAAGCAACCGGG 360

QY 379 GACATCTATGCTATGAAGTGATGAAGAAGAGCTTTATTGGCCAGGAGCAGGTTTCA 438

Db 361 GACATCTATGCTATGAAGTGATGAAGAAGAGCTTTATTGGCCAGGAGCAGGTTTCA 420

QY 439 TTTTTCGAGGAAGCGGAACATATATTCTCGAAGCACAAGCCGCTGGATCCCCCAATTA 498

Db 421 TTTTTCGAGGAAGCGGAACATATATTCTCGAAGCACAAGCCGCTGGATCCCCCAATTA 480

QY 499 CAGTATGCTTTCAGGACAAAATCACCCTTATCTGATGGAGGAATATCAGCTCGAGGG 558

Db 481 CAGTATGCTTTCAGGACAAAATCACCCTTATCTGATGGAGGAATATCAGCTCGAGGG 540

QY 559 GACTTGTGTCACCTTTTGAATAGATATGAGACCAAGTTAGATGAAAACCTGATACAGTTT 618

Db 541 GACTTGTGTCACCTTTTGAATAGATATGAGACCAAGTTAGATGAAAACCTGATACAGTTT 600

QY 619 TACCTAGCTGAGCTGATTTGGCTGTTTCACAGGCTTCATCTGATGGATACGTCGATCGA 678

Db 601 TACCTAGCTGAGCTGATTTGGCTGTTTCACAGGCTTCATCTGATGGATACGTCGATCGA 660

QY 679 GACATCAAGCTGAGAACTTCCTCGTTGACCGCAGGACACACATCAAGCTGGTGGATTTT 738

Db 661 GACATCAAGCTGAGAACTTCCTCGTTGACCGCAGGACACACATCAAGCTGGTGGATTTT 720

QY 739 GGATCTCCCGGAAATGAATTAACAAGATGGTGAATCCCAAACTCCCGATGGGACC 798

Db 721 GGATCTCCCGGAAATGAATTAACAAGATGGTGAATCCCAAACTCCCGATGGGACC 780

QY 799 CCAGATTACATGCTCCTCGTGAAGTGTGACTGATGAACGGGGATGAAAAGGACCTAC 858

Db 781 CCAGATTACATGCTCCTCGTGAAGTGTGACTGATGAACGGGGATGAAAAGGACCTAC 840

QY 859 GGCTGGACTGTGACTGGTGGTTCAGTGGGCGTGAATTCCTATGAGATGATTTATGGAGA 918

Db 841 GGCTGGACTGTGACTGGTGGTTCAGTGGGCGTGAATTCCTATGAGATGATTTATGGAGA 900

QY 919 TCCCTCTCCAGAGGAACTCTGCCAGAACCTTCAATAACATTATGATTTCCAGCGG 978

Db 901 TCCCTCTCCAGAGGAACTCTGCCAGAACCTTCAATAACATTATGATTTCCAGCGG 960

QY 979 TTTTGTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGTATCTGATTTCAAAGC 1038

Db 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGTATCTGATTTCAAAGC 1020

QY 1039 TTGTTGTGGCCGAGAGAGACTGAGTTGAAGGCTTTGCTGCCATCCTTCTTC 1098

Db 1021 TTGTTGTGCGCCAGAAAGAGAGACTGAAGTTGAAGTCTTTGTGTCATCCTTTCTTC 1080  
Qy TCTAAATTTGACTGGAAACAACATTCGTAACTCTCTCTCCCTTCGTTCCTCCACCTCAAG 1158  
Db TCTAAATTTGACTGGAAACAACATTCGTAACTCTCTCTCCCTTCGTTCCTCCACCTCAAG 1140  
Qy TCTGACGATGACACCTCCAAATTTGATGAAACAGAGAAAGATTCGTGGGTTTCATCTCT 1218  
Db TCTGACGATGACACCTCCAAATTTGATGAAACAGAGAAAGATTCGTGGGTTTCATCTCT 1200  
Qy CCGTGCAGCTGAGCCCTCAGGCTCTCGGGTGAAGAACTCCCGTTTGTGGGTTTTTCG 1278  
Db CCGTGCAGCTGAGCCCTCAGGCTCTCGGGTGAAGAACTCCCGTTTGTGGGTTTTTCG 1260  
Qy TACAGCAAGGCACTGGGATTTCTGGTAGATCTGAGTCTGTGTGTCTCGGGTCTGACTCC 1338  
Db TACAGCAAGGCACTGGGATTTCTGGTAGATCTGAGTCTGTGTGTCTCGGGTCTGACTCC 1320  
Qy CCGTGCAGCTGAGCTCCATGAGAAAGAACTTCATCAAAAGCAAGAGCTACAAGAC 1398  
Db CCGTGCAGCTGAGCTCCATGAGAAAGAACTTCATCAAAAGCAAGAGCTACAAGAC 1380  
Qy TCTCAGGCAAGTGTCAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGCA 1458  
Db TCTCAGGCAAGTGTCAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGCA 1440  
Qy GAGGTGGAGCTGTGCTTAGTCAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1518  
Db GAGGTGGAGCTGTGCTTAGTCAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1500  
Qy TCCCTCTCGAGCAGGACCTTCTACCTACATCACAGATGCAAGTCTTAAAGCGAAGT 1578  
Db TCCCTCTCGAGCAGGACCTTCTACCTACATCACAGATGCAAGTCTTAAAGCGAAGT 1560  
Qy TTGGAGCAAGCGAGTGTGCTGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1638  
Db TTGGAGCAAGCGAGTGTGCTGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620  
Qy GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGAGTACAGGCTCAA 1698  
Db GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGAGTACAGGCTCAA 1680  
Qy GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758  
Db GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
Qy CCGAGTGTCTCTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1818  
Db CCGAGTGTCTCTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1800  
Qy CGGAAGCCACAGAAATGTCAGATAACTGTTCGAAGGCTTAAGGATCAAGGGAAGCTGAA 1878  
Db CGGAAGCCACAGAAATGTCAGATAACTGTTCGAAGGCTTAAGGATCAAGGGAAGCTGAA 1860  
Qy GTGGGAGAAATGCGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938  
Db GTGGGAGAAATGCGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
Qy CTTCAAGAGAACTGGAGAGGCT----- 1962  
Db CTTCAAGAGAACTGGAGAGGCTGTAAAGCCAGCAGGAGCCACCGAGCTGCTGCAG 1980  
Qy -----GCAAGGAGCGAGCCGAGAGGAGCTGGAGAGCTGAGAGCCGAGAG 2010  
Db AATATCCGCCAGGCAAGGAGCGAGCCGAGAGGAGCTGGAGAGCTGAGAGCCGAGAG 2040  
Qy GATTCTTCTGAAGGATCAAGAGAGCTGGTGGAGCTGAGAGCGCCGCTTCTCTG 2070  
Db GATTCTTCTGAAGGATCAAGAGAGCTGGTGGAGCTGAGAGCGCCGCTTCTCTG 2100  
Qy GAGAACAAAGTAAAGAGCTAGAGACCATGGAGCGTAGAGAAACAGACTGAAGGATGAC 2130  
Db GAGAACAAAGTAAAGAGCTAGAGACCATGGAGCGTAGAGAAACAGACTGAAGGATGAC 2160

Qy 2131 ATCCAGACAAAAATCCCAACAGATCCAGCAGATGGCTGATAAAAATTTGGAGCTCGAAGAG 2190  
Db ATCCAGACAAAAATCCCAACAGATCCAGCAGATGGCTGATAAAAATTTGGAGCTCGAAGAG 2220  
Qy AARACATCGGAGGCCCAAGTCTCAGCCACAGCTCAGAGTGACCTGAAACAGAAAGAG 2250  
Db AARACATCGGAGGCCCAAGTCTCAGCCACAGCTCAGAGTGACCTGAAACAGAAAGAG 2280  
Qy CAGCACTATGAGGAAAAAGATTTAAAGTGTGGACAAATCAGATAAAGAAACCTTGGCTGAC 2310  
Db CAGCACTATGAGGAAAAAGATTTAAAGTGTGGACAAATCAGATAAAGAAACCTTGGCTGAC 2340  
Qy AAGGAGACATCGAGAACATGATCGAGAGACAGAGGAGGAGCCCATGAGAGGGGCAAA 2370  
Db AAGGAGACATCGAGAACATGATCGAGAGACAGAGGAGGAGCCCATGAGAGGGGCAAA 2400  
Qy ATTCTCAGCAACAGAGGCGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2430  
Db ATTCTCAGCAACAGAGGCGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2460  
Qy CAGAGGATTTGTGAACTGTCTGAAGCCCAATTAAGTCTGAGGAGGAGGAGGAGGAGG 2490  
Db CAGAGGATTTGTGAACTGTCTGAAGCCCAATTAAGTCTGAGGAGGAGGAGGAGGAGG 2520  
Qy CAAGGAAACATGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2550  
Db CAAGGAAACATGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580  
Qy CTGAGACACAGGCTGGGAAAGTTGGAGGCCAGAACCGAAACTGGAGGAGAGCTGGAG 2610  
Db CTGAGACACAGGCTGGGAAAGTTGGAGGCCAGAACCGAAACTGGAGGAGAGCTGGAG 2640  
Qy AAGATCAGCCACCAAGACCAAGTGAAGAAATCGGCTGCTGGAACCTGGAGAGCAAGATTG 2670  
Db AAGATCAGCCACCAAGACCAAGTGAAGAAATCGGCTGCTGGAACCTGGAGAGCAAGATTG 2700  
Qy CCGGAGTCTCAGTCTAGAGCAGGAGGAGCAAACTGGAGCTCAAGCCGCTCAGCTCACAG 2730  
Db CCGGAGTCTCAGTCTAGAGCAGGAGGAGCAAACTGGAGCTCAAGCCGCTCAGCTCACAG 2760  
Qy CTAAGCTCTCCTGAGGAGCGGAGTCAAGTTGACAGCTTGAAGCTCAAGCCGCTCAGCGG 2790  
Db CTAAGCTCTCCTGAGGAGCGGAGTCAAGTTGACAGCTTGAAGCTCAAGCCGCTCAGCGG 2820  
Qy GCCCTGAGAGCGGAGTTCGCCAGGCAAGACAGAGCTGGAAGAGACCAAGAGAGCT 2850  
Db GCCCTGAGAGCGGAGTTCGCCAGGCAAGACAGAGCTGGAAGAGACCAAGAGAGCT 2880  
Qy GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGCCAAATTTGATGCT 2910  
Db GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGCCAAATTTGATGCT 2940  
Qy CTTCTAACAGCTGTACTGTAAATCAGAGCTGGAGGAGCAGCTTAACCCAGCTGAGCCGAG 2970  
Db CTTCTAACAGCTGTACTGTAAATCAGAGCTGGAGGAGCAGCTTAACCCAGCTGAGCCGAG 3000  
Qy GACAACTGAACTCAACCAACCAAACTTCTACTTTGTCAAAACAACTCGATGAGGCTTCT 3030  
Db GACAACTGAACTCAACCAACCAAACTTCTACTTTGTCAAAACAACTCGATGAGGCTTCT 3060  
Qy GGCCCAACAGAGATTTACAACTCGAAGTGAAGTGAACCTTCCCGCGGAGATC 3090  
Db GGCCCAACAGAGATTTACAACTCGAAGTGAAGTGAACCTTCCCGCGGAGATC 3120  
Qy ACGGAAACAGAGATGAGCTTACCGAGCCAGAAACAGATGAGGCTCTGAAGACCAAG 3150  
Db ACGGAAACAGAGATGAGCTTACCGAGCCAGAAACAGATGAGGCTCTGAAGACCAAG 3180  
Qy TGCACCATGCTGGAGGAAACAGGCTATGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3210  
Db TGCACCATGCTGGAGGAAACAGGCTATGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3240

QY 3211 AAAGAGCGGAGTGGGAGGCTCGAGAGCGTCTCGTGATGAGAAATCCAGTTTGAG 3270  
DB 3241 AAAGAGCGGAGTGGGAGGCTCGAGAGCGTCTCGTGATGAGAAATCCAGTTTGAG 3300  
QY 3271 TGTGGGTTCCGAGAGTGCAGAGAAATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3330  
DB 3301 TGTGGGTTCCGAGAGTGCAGAGAAATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3360  
QY 3331 GATCAGCGGATCACCGAGTCTCCGAGTGGTGGAGCTGGCAGTGAAGGACAGAGGCT 3390  
DB 3361 GATCAGCGGATCACCGAGTCTCCGAGTGGTGGAGCTGGCAGTGAAGGACAGAGGCT 3420  
QY 3391 GAGATTCTCGTCTGCAGCAGGCTCTCAAGAGCAGAAAGCTGAAGCCGAGAGCCTCTCT 3450  
DB 3421 GAGATTCTCGTCTGCAGCAGGCTCTCAAGAGCAGAAAGCTGAAGCCGAGAGCCTCTCT 3480  
QY 3451 GACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCGAGACTTA 3510  
DB 3481 GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCGAGACTTA 3540  
QY 3511 CAGCAGAAAGCTGAGAGTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3570  
DB 3541 CAGCAGAAAGCTGAGAGTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3600  
QY 3571 TTACAGCAGAGATGGACCTGCAGAAATCAATTTCCGTCGACTCAAGAGCTGCAA 3630  
DB 3601 TTACAGCAGAGATGGACCTGCAGAAATCAATTTCCGTCGACTCAAGAGCTGCAA 3660  
QY 3631 GAAGCTCTAGATCGGCTGATCTACTGAACAGAGAGAGTGAAGTGAAGTATCAGCTG 3690  
DB 3661 GAAGCTCTAGATCGGCTGATCTACTGAACAGAGAGAGTGAAGTGAAGTATCAGCTG 3720  
QY 3691 GAAACATTCAGGTTCTTATCTCATGAAAGGTGAAATGGAAGCACTATTTCTCAA 3750  
DB 3721 GAAACATTCAGGTTCTTATCTCATGAAAGGTGAAATGGAAGCACTATTTCTCAA 3780  
QY 3751 CAACCAAACTCATGATTTCTGCAAGCCAAATGGAACCACTCTTAAAGAAAGAG 3810  
DB 3781 CAACCAAACTCATGATTTCTGCAAGCCAAATGGAACCACTCTTAAAGAAAGAG -- 3838  
QY 3811 GGTTTATTTAGTCGAGGAAAGAGGACCTGCTTTACCCACAGAGCTTCTCTGCAGTAC 3870  
DB 3839 -----AGTTCCTCTGCAGTAC 3855  
QY 3871 AATGAGCTGAAGTGGCCCTGGAGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGCC 3930  
DB 3856 AATGAGCTGAAGTGGCCCTGGAGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGCC 3915  
QY 3931 CTTGAGAAACCCGATCGAGCTCCGGTCCGCCGGGAGGAGAGCTCCCAACCGCAAGCA 3990  
DB 3916 CTTGAGAAACCCGATCGAGCTCCGGTCCGCCGGGAGGAGAGCTCCCAACCGCAAGCA 3975  
QY 3991 ACAGCAACCCCAACCCATCCAGCAGCAACCGCAGGAGCAGATCGCCATGTCGGCC 4050  
DB 3976 ACAGCAACCCCAACCCATCCAGCAGCAACCGCAGGAGCAGATCGCCATGTCGGCC 4035  
QY 4051 ATCGTGGGTCCGACAGACAGCCAGTGCATCAGCTGCTGSCCCGCCATCCAGC 4110  
DB 4036 ATCGTGGGTCCGACAGACAGCCAGTGCATCAGCTGCTGSCCCGCCATCCAGC 4095  
QY 4111 CGCAGAAAGAGTCTTCAACTCCAGAGGAAATTTAGTCGGCTGTTAAGGAAACCGCATGCAC 4170  
DB 4096 CGCAGAAAGAGTCTTCAACTCCAGAGGAAATTTAGTCGGCTGTTAAGGAAACCGCATGCAC 4155  
QY 4171 CACAATATTCCTCACCGATTACGTTAGGACTGAACATCGGAGCCCAAGTGTCTGTG 4230  
DB 4156 CACAATATTCCTCACCGATTACGTTAGGACTGAACATCGGAGCCCAAGTGTCTGTG 4215  
QY 4231 TGTCTGGATACCGTGCATTTTGACGCCAGGATCCAAATGCTTCAAGTGTGAGTGTG 4290  
DB 4216 TGTCTGGATACCGTGCATTTTGACGCCAGGATCCAAATGCTTCAAGTGTGAGTGTG 4275  
QY 4291 TGTCAACCCCAAGTGTCTCCAGCTGTTGCCAGCCACCTGCGGCTTGCTGCTGAAATGCC 4350

DB 4276 TGTCAACCCCAAGTGTCTCCAGTGTCTGCCAGCACCTGCGGCTTGCTGCTGATATGCC 4335  
QY 4351 ACACACTTCAACGAGAGCTTCTGCCGTGACAAATGAATGAATCTCCCGAGTCTCCAGACCAAG 4410  
DB 4336 ACACACTTCAACGAGAGCTTCTGCCGTGACAAATGAATGAATCTCCCGAGTCTCCAGACCAAG 4395  
QY 4411 GAGCCACGACAGCTTGACCTGCAAGAGGTGGATGAAGTGCACAGGAATAACAAACGA 4470  
DB 4396 GAGCCACGACAGCTTGACCTGCAAGAGGTGGATGAAGTGCACAGGAATAACAAACGA 4455  
QY 4471 GGAAGCAAGCTCGGACAGGAAGTACATTTCTGAGGGATCAAAAGTCTCATTTAT 4530  
DB 4456 GGAAGCAAGCTCGGACAGGAAGTACATTTCTGAGGGATCAAAAGTCTCATTTAT 4515  
QY 4531 GACATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCC 4590  
DB 4516 GACATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCC 4575  
QY 4591 GACGGGATGATCTATTTATGCTGGCTTGGTCTTCCGAATCTGCAAAATACAGCCAAA 4650  
DB 4576 GACGGGATGATCTATTTATGCTGGCTTGGTCTTCCGAATCTGCAAAATACAGCCAAA 4635  
QY 4651 GCA----- 4653  
DB 4636 GCAGATGTCCATACATCTGAAAGATGGAATCTCACCCGACACCACTGCTGCGCCGGG 4695  
QY 4654 ----- 4653  
DB 4696 AGAACCTCTACTTGTAGTCTCCAGCTTCCCTGCACAAACAGCGCTGGGTACCGCCTTA 4755  
QY 4654 -----GAAAAAGCAGAGCTGATGCTAACTG 4680  
DB 4756 GAATCAGTTGTGCGAGTGGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAACTG 4815  
QY 4681 CTTGAAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACTGCACGCTG 4740  
DB 4816 CTTGAAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACTGCACGCTG 4875  
QY 4741 CCCTTCAGTACACAGGTGGTGTGGTGGGCAACGAGAAAGGCTCTACGCCCTGAATGTC 4800  
DB 4876 CCCTTCAGTACACAGGTGGTGTGGTGGGCAACGAGAAAGGCTCTACGCCCTGAATGTC 4935  
QY 4801 TTGAAAACTCCCTAAACCATGTCCTCCAGGAATTTGGAGCAGTCTTCAAAATTTATATTC 4860  
DB 4936 TTGAAAACTCCCTAAACCATGTCCTCCAGGAATTTGGAGCAGTCTTCAAAATTTATATTC 4995  
QY 4861 AAGGACCTGGAGAGTACTCATGATAGCAGGAAGAGCGGCACTGTCTTTGTGGAC 4920  
DB 4996 AAGGACCTGGAGAGTACTCATGATAGCAGGAAGAGCGGCACTGTCTTTGTGGAC 5055  
QY 4921 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 4980  
DB 5056 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 5115  
QY 4981 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTGTTTGGGGCAGCAGAAAGTTGAAAC 5040  
DB 5116 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTGTTTGGGGCAGCAGAAAGTTGAAAC 5175  
QY 5041 GGGCTCTGCATCTGTGAGCCATGCCAGCAAGTGTGCTATTTCTCCGCTACAGAAAC 5100  
DB 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAGTGTGCTATTTCTCCGCTACAGAAAC 5235  
QY 5101 CTCAGCAAAATCTCATCTCCGAAAGAGATAGACCTTCAGAGCCCTGCAGCTGTATCCAC 5160  
DB 5236 CTCAGCAAAATCTCATCTCCGAAAGAGATAGACCTTCAGAGCCCTGCAGCTGTATCCAC 5295  
QY 5161 TTCAACCAATTTACAGTATCTCTCATTTGGAACCAATAAATTTACGAAATTCGACATGAAGCAG 5220  
DB 5296 TTCAACCAATTTACAGTATCTCTCATTTGGAACCAATAAATTTACGAAATTCGACATGAAGCAG 5355  
QY 5221 TACAGCTCGAGGAATTCCTGGATGAAGTGAACATTTCTTGGCAGCTGTGTTTGGC 5280

Db 5356 TACACGCTCGAGGATTCCTGGATAGAATGACCATTCCTTGGCACCTCTGTGTTGCC 5415  
QY 5281 GCCTCTTCCAAAGCTTCCCTGCTCAATCGTGCAGTGAACAGCGCAGGGCAGGAGAG 5340  
Db 5416 GCCTCTTCCAAAGCTTCCCTGCTCAATCGTGCAGTGAACAGCGCAGGGCAGGAGAG 5475  
QY 5341 GAGTACTGCTGTGTTTCCAGAAATTGAGTGTTCGTCGATTTTACGGAAGACGTAGC 5400  
Db 5476 GAGTACTGCTGTGTTTCCAGAAATTGAGTGTTCGTCGATTTTACGGAAGACGTAGC 5535  
QY 5401 CGCACAGAGACTCAAGTGGAGTGGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT 5460  
Db 5536 CGCACAGAGACTCAAGTGGAGTGGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT 5595  
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACGCTCCTCAGCA 5520  
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGACGCTCCTCAGCA 5655  
QY 5521 GGGACCCCTGCCGAGCGTACCTGGACATCCCGAACCCCGCTACCTGGGCCCTGCCATT 5580  
Db 5656 GGGACCCCTGCCGAGCGTACCTGGACATCCCGAACCCCGCTACCTGGGCCCTGCCATT 5715  
QY 5581 TCCTCAGGAGCGATTTACTTTGGCTCTCATACAGGATAAATTAAAGGTCATTTTGTCTGC 5640  
Db 5716 TCCTCAGGAGCGATTTACTTTGGCTCTCATACAGGATAAATTAAAGGTCATTTTGTCTGC 5775  
QY 5641 AAGGGAACCTCGTGAAGAGTCGGCACTGAACACACCGGGCGCGTCCACCTCCCGC 5700  
Db 5776 AAGGGAACCTCGTGAAGAGTCGGCACTGAACACACCGGGCGCGTCCACCTCCCGC 5835  
QY 5701 AGCAGCCCCCAAGAGGAGCGCCACCCAGCTACACAGACATCACCAAGCGCGTGCC 5760  
Db 5836 AGCAGCCCCCAAGAGGAGCGCCACCCAGCTACACAGACATCACCAAGCGCGTGCC 5895  
QY 5761 TCAGCCCCAGCGCGCCGAGGCGCCAGCCACCCGCGAGAGCCCAAGCACACCCACCGC 5820  
Db 5896 TCAGCCCCAGCGCGCCGAGGCGCCAGCCACCCGCGAGAGCCCAAGCACACCCACCGC 5955  
QY 5821 TACCGGAGGCGGAGCGAGCTGCGCAGGACAGTCTCTGGCGGCCCTGAGAGCA 5880  
Db 5956 TACCGGAGGCGGAGCGAGCTGCGCAGGACAGTCTCTGGCGGCCCTGAGAGCA 6015  
QY 5881 GAGAACTCCCCCGGCGGATGCTCAGACCGCGAGAGCGGTGCCCGGAGGCTGTT 5940  
Db 6016 GAGAACTCCCCCGGCGGATGCTCAGACCGCGAGAGCGGTGCCCGGAGGCTGTT 6075  
QY 5941 GAAGACAGCAGAGGGGCGCGTGCCTGCGGAGCGGTGAGGACCCCGTGTCCAGGTG 6000  
Db 6076 GAAGACAGCAGAGGGGCGCGTGCCTGCGGAGCGGTGAGGACCCCGTGTCCAGGTG 6135  
QY 6001 AACAGGGGAGAGGGCA 6017  
Db 6136 AACAGGTGAGGAGCA 6152

RESULT 5  
AAL55214  
ID AAL55214 standard; DNA; 6165 BP.  
XX AC AAL55214;  
XX DT  
XX DT 01-MAY-2003 (first entry)  
XX DE Human CR1K encoding DNA sequence, SEQ ID NO 1.  
XX KW Anorectic; hypotensive; cardiatic; antilipaeamic; cerebroprotective;  
KW antitumor; osteopathic; antiarthritic; cystostatic; antidepressant;  
KW immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;  
KW neuroprotective; antiinflammatory; antidiabetic; analgesic;  
KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;  
KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;  
KW central nervous system disorder; chronic obstructive pulmonary disease;  
KW diabetes; pain; gene; ds.

XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH CDS 1..6165  
FT /\*tag= a  
FT /product= "Human CR1K protein"  
XX  
XX WO2003004523-A1.  
XX  
XX 16-JAN-2003.  
XX  
XX 28-JUN-2002; 2002WO-EP007156.  
XX  
XX 02-JUL-2001; 2001US-0301841P.  
XX 11-DEC-2001; 2001US-0338651P.  
XX 25-APR-2002; 2002US-0375014P.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Zhu Z;  
XX WPI; 2003-221576/21.  
XX P-PSDB; AAO26959.  
XX  
XX New human citron rho/rac-interacting kinase (CR1K) polypeptide and  
PT polynucleotide, useful in preventing, ameliorating or treating diseases  
PT associated with human CR1K dysfunction, e.g. obesity, diabetes or  
PT Alzheimer's disease.  
XX  
XX Example 1; Fig 1; 237pp; English.  
XX  
XX The invention relates to an isolated polynucleotide encoding a human  
CC citron rho/rac-interacting kinase polypeptide. The isolated  
CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the  
CC specification. The human citron rho/rac-interacting kinase (CR1K)  
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or  
CC treating diseases associated with human CR1K dysfunction such as obesity  
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery  
CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of  
CC cancer including endometrial, breast, prostate and colon cancer),  
CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood  
CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's  
CC disease), chronic obstructive pulmonary disease, or diabetes. These can  
CC also be used to treat pain associated with the disorders. The human CR1K  
CC polypeptide is also useful in diagnostic assays or in genetic testing.  
CC The expression vector or the reagent is useful in preparing a medicament  
CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a  
CC central nervous system disorder, or chronic obstructive pulmonary  
CC disease. The fusion protein is useful for generating antibodies against a  
CC CR1K polypeptide and for use in various assay systems. The methods are  
CC useful in producing and detecting the polynucleotide and polypeptide and  
CC in screening for agents that modulate the activity of the human CR1K  
CC polypeptide. This polynucleotide sequence represents a DNA sequence  
CC encoding a human CR1K protein of the invention  
XX  
XX Sequence 6165 BP; 1735 A; 1549 C; 1680 G; 1201 T; 0 U; 0 Other;  
Query Match 86.2%; Score 5666.2; DB 8; Length 6165;  
Best Local Similarity 95.9%; Pred.No. 0;  
Matches 5947; Conservative 0; Mismatches 13; Indels 243; Gaps 3;  
QY 19 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTTGGATGCTGCTGTGACCCATT 78  
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTTGGATGCTGCTGTGACCCATT 60  
QY 79 GCCAGCGGGCTCCAGGCTGAATCTTTCTTCAGGGGAAACACCCCTTATGACTCAA 138  
Db 61 GCCAGCGGGCTCCAGGCTGAATCTTTCTTCAGGGGAAACACCCCTTATGACTCAA 120  
QY 139 CAGCAGATGCTCTCTCTTCCCGAGAGGATATTAGATCCCTCTTTGTTCTCTTTGAA 198  
Db 121 CAGCAGATGCTCTCTCTTCCCGAGAGGATATTAGATCCCTCTTTGTTCTCTTTGAA 180



Db 2341 |||||AAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGAGGCCCATGAGAAAGGCGAAA||| 2400  
QY 2371 ATTCTCAGCGAAACAGAGGCGGATGATCAATGTGTATGGATTCCAAGATCAGATCCCTGGAA||| 2430  
Db 2401 ATTCTCAGCGAAACAGAGGCGGATGATCAATGTGTATGGATTCCAAGATCAGATCCCTGGAA||| 2460  
QY 2431 CAGAGAGTTGTGGAACTGTCTGAAGCCATAAATCTGAGCAAAATAGCAGTCTTTTACC||| 2490  
Db 2461 CAGAGAGTTGTGGAACTGTCTGAAGCCATAAATCTGAGCAAAATAGCAGTCTTTTACC||| 2520  
QY 2491 CAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAAAATTTTAC||| 2550  
Db 2521 CAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAAAATTTTAC||| 2580  
QY 2551 CTGGAGACACAGCTGGGAAGTTGGAGGCCAGAACCGGAAAACTGGAGAGCAGCTGGAG||| 2610  
Db 2581 CTGGAGACACAGCTGGGAAGTTGGAGGCCAGAACCGGAAAACTGGAGAGCAGCTGGAG||| 2640  
QY 2611 AAGATCAGCCACCAAGACACACAGTGCAGAGATCGCTGCTGAACTGGAGTGGAGCAAGATTG||| 2670  
Db 2641 AAGATCAGCCACCAAGACACACAGTGCAGAGATCGCTGCTGAACTGGAGTGGAGCAAGATTG||| 2700  
QY 2671 CGGAGGTCAGTCTAGAGCACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAG||| 2730  
Db 2701 CGGAGGTCAGTCTAGAGCACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCACAG||| 2760  
QY 2731 CTACAGCTCTCCCTGCAGAGCGGAGTCACAGTTGACAGCCCTCAGGCTGCACGGGG||| 2790  
Db 2761 CTACAGCTCTCCCTGCAGAGCGGAGTCACAGTTGACAGCCCTCAGGCTGCACGGGG||| 2820  
QY 2791 GGCCTGGAGAGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGCAAGCT||| 2850  
Db 2821 GGCCTGGAGAGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGCAAGCT||| 2880  
QY 2851 GAAGAGGATCCAGCAGCTACCGCACATAGATGAAATCCAGGCGAAATTTGATGCT||| 2910  
Db 2881 GAAGAGGATCCAGCAGCTACCGCACATAGATGAAATCCAGGCGAAATTTGATGCT||| 2940  
QY 2911 CTTCTGAAGCTGTACTGTATACAGACTCTGGAGGACAGCTAAACAGCTGACCCGAG||| 2970  
Db 2941 CTTCTGAAGCTGTACTGTATACAGACTCTGGAGGACAGCTAAACAGCTGACCCGAG||| 3000  
QY 2971 GACACGCTGAACTCAACACCAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCT||| 3030  
Db 3001 GACACGCTGAACTCAACACCAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCT||| 3060  
QY 3031 GCGGCCAAGCAGAGATTGTAACAATCGGAAGTGAAGTGAACATCTCCGCGGAGATC||| 3090  
Db 3061 GCGGCCAAGCAGAGATTGTAACAATCGGAAGTGAAGTGAACATCTCCGCGGAGATC||| 3120  
QY 3091 ACGGAACGAGATGAGCTTACCGCCAGAGCAAAACGATGAGGCTCTGAAGACCACG||| 3150  
Db 3121 ACGGAACGAGATGAGCTTACCGCCAGAGCAAAACGATGAGGCTCTGAAGACCACG||| 3180  
QY 3151 TGCACCATCTGGAGGAAACAGTATGGAATTTGGAGGCCCTAAAACGATGCTGTAGAA||| 3210  
Db 3181 TGCACCATCTGGAGGAAACAGTATGGAATTTGGAGGCCCTAAAACGATGCTGTAGAA||| 3240  
QY 3211 AAAGAGCGCAGTGGGAGCGCTGGAGAGCGTCTCGGCTGATGAGAAATCCCAAGTTTGA||| 3270  
Db 3241 AAAGAGCGCAGTGGGAGCGCTGGAGAGCGTCTCGGCTGATGAGAAATCCCAAGTTTGA||| 3300  
QY 3271 TGTTCGGGTTGAGAGTGCAGAGAAATGCTGGACACCGAGAAAACAGAGCGGCGAGAGCC||| 3330  
Db 3301 TGTTCGGGTTGAGAGTGCAGAGAAATGCTGGACACCGAGAAAACAGAGCGGCGAGAGCC||| 3360  
QY 3331 GATCAGGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGAGTGGAGTGAAGGACACAGGCT||| 3390  
Db 3361 GATCAGGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGAGTGGAGTGAAGGACACAGGCT||| 3420  
QY 3391 GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGAGCAGAAAGCTGAAGCCGAGAGCCTCTCT||| 3450

Db 3421 GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGACGAGAGCTGAAGCCGAGAGCCTCTCT||| 3480  
QY 3451 GACAAGCTCAATGACCTGGAGAAAGACATGCTATGTTGAAATGAATGCCCGAAGCTTA||| 3510  
Db 3481 GACAAGCTCAATGACCTGGAGAAAGACATGCTATGTTGAAATGAATGCCCGAAGCTTA||| 3540  
QY 3511 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAGACCAAA||| 3570  
Db 3541 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAGACCAAA||| 3600  
QY 3571 TTACAGCAGCAGATGGACCTGCAGAAAAATCACATTTTCCGCTCTGACTCAAGGACTCAA||| 3630  
Db 3601 TTACAGCAGCAGATGGACCTGCAGAAAAATCACATTTTCCGCTCTGACTCAAGGACTCAA||| 3660  
QY 3631 GAACTCTAGATCGGCTGATCTACTGAACACAGAAAGAGTGAATTTTCTTCTCAA||| 3690  
Db 3661 GAACTCTAGATCGGCTGATCTACTGAACACAGAAAGAGTGAATTTTCTTCTCAA||| 3720  
QY 3691 GAAAAACATTCAAGGTTCTTATTTCTATGAAAAAGTGAATAATGGAAGGCACTATTTCTCAA||| 3750  
Db 3721 GAAAAACATTCAAGGTTCTTATTTCTATGAAAAAGTGAATAATGGAAGGCACTATTTCTCAA||| 3780  
QY 3751 CAAACCAAACTCAATGATTTTCTGCAAGCCAAATGGAACCAACTGCTAAAAAGAAAAAG||| 3810  
Db 3781 CAAACCAAACTCAATGATTTTCTGCAAGCCAAATGGAACCAACTGCTAAAAAGAAAA--||| 3838  
QY 3811 GGTTTATTAGTCGACGAAAAAGGAGCCCTGCTTTACCCACACAGGTTCTCTCAGTAC||| 3870  
Db 3839 -----AGTTCTCTGCAGTAC||| 3855  
QY 3871 AATCAGCTGAAGCTGGCCCTGGAGAAAGAAAGCTCGTGTGCAGAGCTAGAGAAAGCC||| 3930  
Db 3856 AATCAGCTGAAGCTGGCCCTGGAGAAAGAAAGCTCGTGTGCAGAGCTAGAGAAAGCC||| 3915  
QY 3931 CTTCAGAGACCCGATCGAGCTCCGCTCCGCGGAGGAAGCTGCCACCGCAAGCA||| 3990  
Db 3916 CTTCAGAGACCCGATCGAGCTCCGCTCCGCGGAGGAAGCTGCCACCGCAAGCA||| 3975  
QY 3991 ACGGACCAACCCACACCCATCCAGCCAGCCACCGCGAGGAGCAGATCGCCATGTCGCC||| 4050  
Db 3976 ACGGACCAACCCACACCCATCCAGCCAGCCACCGCGAGGAGCAGATCGCCATGTCGCC||| 4035  
QY 4051 ATCGTGGTGCAGAGACACAGCCAGTGCATGAGCTGTGCGCCCGCCCTCCAGC||| 4110  
Db 4036 ATCGTGGTGCAGAGACACAGCCAGTGCATGAGCTGTGCGCCCGCCCTCCAGC||| 4095  
QY 4111 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTTTAAGGAAACGATGCAC||| 4170  
Db 4096 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTTTAAGGAAACGATGCAC||| 4155  
QY 4171 CACAATATTCCTCAACGATTCAACGTAGGACTGAACATGCGAGCCACAAAGTGTGTGTG||| 4230  
Db 4156 CACAATATTCCTCAACGATTCAACGTAGGACTGAACATGCGAGCCACAAAGTGTGTGTG||| 4215  
QY 4231 TGTCTGGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGCTCTCGAATGTCAGGTGATG||| 4290  
Db 4216 TGTCTGGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGCTCTCGAATGTCAGGTGATG||| 4275  
QY 4291 TGTCAACCCCAAGTCTCCAGTGTCTTCCAGCCACCTGCGGCTTGCCTGTGATATGCC||| 4350  
Db 4276 TGTCAACCCCAAGTCTCCAGTGTCTTCCAGCCACCTGCGGCTTGCCTGTGATATGCC||| 4335  
QY 4351 ACACATTTACCGAGGCTTCTGCCGTGACAAAATGAACTCCCGAGGCTTCCAGACCAAG||| 4410  
Db 4336 ACACATTTACCGAGGCTTCTGCCGTGACAAAATGAACTCCCGAGGCTTCCAGACCAAG||| 4395  
QY 4411 GAGCCACGACGACGCTTTCACCTGGAAGGTTGATGAAGTGGCCAGGAATAACAAACGA||| 4470  
Db 4396 GAGCCACGACGACGCTTTCACCTGGAAGGTTGATGAAGTGGCCAGGAATAACAAACGA||| 4455  
QY 4471 GGAAGCAAGGCTGGGACAGAAATCAATTTCTCTGAGGGAATCAAAAGTCTCTCATTTAT||| 4530  
Db 4456 GGAAGCAAGGCTGGGACAGAAATCAATTTCTCTGAGGGAATCAAAAGTCTCTCATTTAT||| 4515

```
QY 4531 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCC 4590
Db 4516 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCTTCCC 4575
QY 4591 GACGGGATGTATCTATTTCATGTGTGCGTGGTGGTTCCTCCAACTCGCAATACAGCCAAA 4650
Db 4576 GACGGGATGTATCTATTTCATGTGTGCGTGGTGGTTCCTCCAACTCGCAATACAGCCAAA 4635
QY 4651 GCA----- 4653
Db 4636 GCAGATGCCATACATACACTGAAGATGGAATCTCAOCCGGCACACCACTGCTGGCCCGGG 4695
QY 4654 ----- 4653
Db 4696 AGAACCCCTCTACTTGTAGTCCAGCTCCAGCTCCCTGACAAAAGCGCTGGGTCAACCGCTTA 4755
QY 4654 -----GAAAAGCAGAACTGATGCTAACTG 4680
Db 4756 GAATCAGTTGTGCGAGGTGGAGAGTTTCTAGGGAAGAAAGCAGAACTGATGCTAACTG 4815
QY 4681 CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACCTGACGCTG 4740
Db 4816 CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACCTGACGCTG 4875
QY 4741 CCCTTCAGTGAACAGAGTGTGTTGGTGGGCACCGAGGAAGGGCTCTACGCCCTGAATGTC 4800
Db 4876 CCCTTCAGTGACAGAGTGTGTTGGTGGGCACCGAGGAAGGGCTCTACGCCCTGAATGTC 4935
QY 4801 TTGAAAACCTCCCTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAATTTATATTTC 4860
Db 4936 TTGAAAACCTCCCTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAATTTATATTTC 4995
QY 4861 AAGGACCTGGAGAGCTACTCATATAGCAGAGAGAGAGCGGCACTGTCTCTGTGGAC 4920
Db 4996 AAGGACCTGGAGAGCTACTCATATAGCAGAGAGAGAGCGGCACTGTCTCTGTGGAC 5055
QY 4921 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCTGCCAGCCGACATCTCA 4980
Db 5056 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCTGCCAGCCGACATCTCA 5115
QY 4981 CCCAAATTTTGAAGCTGTCAAGGGCTGCCATTTGTTGGGCGAGCAAGATTGAGAAC 5040
Db 5116 CCCAAATTTTGAAGCTGTCAAGGGCTGCCATTTGTTGGGCGAGCAAGATTGAGAAC 5175
QY 5041 GGGCTCTGCAATCTGTGAGCCATGCCAGCAAGTCTGTCATTTCTCGCTACACGAAAC 5100
Db 5176 GGGCTCTGCAATCTGTGAGCCATGCCAGCAAGTCTGTCATTTCTCGCTACACGAAAC 5235
QY 5101 CTCAGCAAACTACTGCAATCGGAAAGATAGAGACCTCAGAGCCCTGACGCTGTATCCAC 5160
Db 5236 CTCAGCAAACTACTGCAATCGGAAAGATAGAGACCTCAGAGCCCTGACGCTGTATCCAC 5295
QY 5161 TTCACCAATPACAGTATCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5220
Db 5296 TTCACCAATPACAGTATCTCATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5355
QY 5221 TACACCTCGAGGAATTCCTGGATAGATGACCATTCCTTGGCACCTGCTGTGTTGCC 5280
Db 5356 TACACCTCGAGGAATTCCTGGATAGATGACCATTCCTTGGCACCTGCTGTGTTGCC 5415
QY 5281 GCCTTCTCAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGAGCGGCGAGGAGAG 5340
Db 5416 GCCTTCTCAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGAGCGGCGAGGAGAG 5475
QY 5341 GAGTACTTGTGTTTCCAGCAATTTGGAGTGTTCGTGGATTTCTACGAAAGCTGAGC 5400
Db 5476 GAGTACTTGTGTTTCCAGCAATTTGGAGTGTTCGTGGATTTCTACGAAAGCTGAGC 5535
QY 5401 CGCACAGAGATCTCAAGTGGAGTCCCTTACCTTTGGCCCTTTGGCTTACAGAGAACCTAT 5460
Db 5536 CGCACAGAGATCTCAAGTGGAGTCCCTTACCTTTGGCCCTTTGGCTTACAGAGAACCTAT 5595
```

```
QY 5461 CTGTTTGTGACCCCACTTCACTCACTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5520
Db 5596 CTGTTTGTGACCCCACTTCACTCACTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5655
QY 5521 GGGACCCCTGCGGAGCGTACTCTGACATCCGAAACCGCGCTACTCTGGGCCCTTGCCTT 5580
Db 5656 GGGACCCCTGCGGAGCGTACTCTGACATCCGAAACCGCGCTACTCTGGGCCCTTGCCTT 5715
QY 5581 TCCTCAGAGCGATTACTTGGCGTCTCTACACAGATAAATTAAGGTCATTGCTGC 5640
Db 5716 TCCTCAGAGCGATTACTTGGCGTCTCTACACAGATAAATTAAGGTCATTGCTGC 5775
QY 5641 AAGGAAACCTCTGTAAGAGTCCGGCACTGAACACACCGGGGCCGCTCCACCTCCGC 5700
Db 5776 AAGGAAACCTCTGTAAGAGTCCGGCACTGAACACACCGGGGCCGCTCCACCTCCGC 5835
QY 5701 AGCAGCCCAACAGAGAGGCCCCACCACTGTAACAGACATCACAAGCGGTGGCC 5760
Db 5836 AGCAGCCCAACAGAGAGGCCCCACCACTGTAACAGACATCACAAGCGGTGGCC 5895
QY 5761 TCCAGCCAGCGCCCGCCGAGAGGCCCCAGCCACCGCGAGAGCAAGCACACCCCAACCGC 5820
Db 5896 TCCAGCCAGCGCCCGCCGAGAGGCCCCAGCCACCGCGAGAGCAAGCACACCCCAACCGC 5955
QY 5821 TACCGAGAGGCGGAGCCGAGCTCGAGTCTCGCAGGACAAGTCTCTGCGCCGCCCTTGAGCGA 5880
Db 5956 TACCGAGAGGCGGAGCCGAGCTCGAGTCTCGCAGGACAAGTCTCTGCGCCGCCCTTGAGCGA 6015
QY 5881 GAGAAGTCCCCCGCCCGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGGCTGTTT 5940
Db 6016 GAGAAGTCCCCCGCCCGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGGCTGTTT 6075
QY 5941 GAACACAGACAGCGGCGCTGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCTCCAGGTG 6000
Db 6076 GAACACAGACAGCGGCGCTGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCTCCAGGTG 6135
QY 6001 AACAGGGAAGAGCGCAGAGTGC 6023
Db 6136 AACAGGTCCTGGACCACTCTTC 6158

RESULT 6
ABQ78870
ID ABQ78870 standard; cDNA; 6165 BP.
AC ABQ78870;
XX
XX 10-OCT-2002 (first entry)
XX
XX Human kinase cDNA #1.
XX
XX Human; kinase; enzyme; serine-threonine kinase; neotropic; cytostatic;
XX Citron rho-interacting kinase; gene therapy; mental disorder; cancer;
XX gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..6165
XX FT /*tag= a
XX FT /product= "Kinase"
XX FT replace(5218,G)
XX FT /*tag= b
XX FT /standard_name= "Single nucleotide polymorphism"
XX FT replace(6065,G)
XX FT /*tag= b
XX FT /standard_name= "Single nucleotide polymorphism"
XX
XX WO200259325-A2.
XX
XX 01-AUG-2002.
XX
XX 20-DEC-2001; 2001WO-US050497.
XX
XX PF
```

XX	27-DEC-2000; 2000US-0258335P.
PR	{LEXI-} LEXICON GENETICS INC.
XX	
XX	Yu X, Miranda M, Friddle CJ;
PI	
XX	WPI; 2002-599796/64.
DR	P-PSDB; ABB81927.
XX	
PT	Novel polynucleotide encoding human proteins that are structurally
PT	similar to animal kinases, useful for drug screening, diagnosis, in gene
PT	therapy of disorders and diseases e.g. cancer and pharmacogenomic
PT	applications.
XX	
PS	Claim 1; Page 37-39; 50pp; English.
XX	
CC	The invention relates to a novel human protein that shares structural
CC	similarity with animal kinases, including serine-threonine kinases,
CC	particularly Citron rho-interacting kinases. The proteins of the
CC	invention have neotropic and cytosolic activity. The polynucleotides may
CC	have a use in gene therapy. The encoded novel polypeptides are useful for
CC	generating antibodies, as reagents in diagnostic assays, for identifying
CC	other cellular gene products related to NHP and as reagents in assays for
CC	screening for compounds that are useful in the treatment of mental,
CC	biological or medical disorders and diseases including cancer. The
CC	sequence encodes a novel human kinase of the invention
XX	
XX	Sequence 6165 BP; 1735 A; 1550 C; 1679 G; 1201 T; 0 U; 0 Other;
SQ	
	Query Match 86.1%; Score 5661.4; DB 6; Length 6165;
	Best Local Similarity 95.8%; Pred. No. 0;
	Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3
QY	19 ATGTTGAAGTTCAAAATATGGACGCGGAATCCITTTGGATGCTGCTGCTGACCCATT 78
DB	1 ATGTTGAAGTTCAAAATATGGACGCGGAATCCITTTGGATGCTGCTGCTGACCCATT 60
QY	79 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCCITTTATGACTCA 138
DB	61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCCITTTATGACTCA 120
QY	139 CAGCAGATGTCCTCTTTCCGAGAGGGATATAGATGCCCTCTTTGTTCTCTTTGAA 198
DB	121 CAGCAGATGTCCTCTTTCCGAGAGGGATATAGATGCCCTCTTTGTTCTCTTTGAA 180
QY	199 GAATGCAGTCAGCCCTGCTCTGATGAAGATTAAACACGTGAGCAACTTTGTCGGAAGTAT 258
DB	181 GAATGCAGTCAGCCCTGCTCTGATGAAGATTAAACACGTGAGCAACTTTGTCGGAAGTAT 240
QY	259 TCCGACACCATAGCTAGTTACAGAGGCTCCAGCTTTCCGGCAAGGACTTTCGAAGTCAGA 318
DB	241 TCCGACACCATAGCTAGTTACAGAGGCTCCAGCTTTCCGGCAAGGACTTTCGAAGTCAGA 300
QY	319 AGTCCTGTAGGTTGGTTCACCTTTGCTCAAGTCAGGTGTTAAGAGAGAAAGCAACCGG 378
DB	301 AGTCCTGTAGGTTGGTTCACCTTTGCTCAAGTCAGGTGTTAAGAGAGAAAGCAACCGG 360
QY	379 GACATCTATGCTATGAAAGTGATGAAGAAAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 438
DB	361 GACATCTATGCTATGAAAGTGATGAAGAAAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 420
QY	439 TTTTTCGAGAGAGCGGNAACATATTATCTCGAGGCAAGCCCGTGGATCCCCCAATT 498
DB	421 TTTTTCGAGAGAGCGGNAACATATTATCTCGAGGCAAGCCCGTGGATCCCCCAATT 480
QY	499 CAGTATGCTCTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 558
DB	481 CAGTATGCTCTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540
QY	559 GACTTCGCTGTACTTTTGAATAGATATGAGGACAGGTTAGATGAAAACTGTATACAGTTT 618
DB	541 GACTTCGCTGTACTTTTGAATAGATATGAGGACAGGTTAGATGAAAACTGTATACAGTTT 600

QY 1699 GTGGAAGAAATGAGTTGATGAATCAGTTGGAAGAGGATCTTGCTCAGCAAGAGA 1758  
DB 1681 GTGGAAGAAATGAGTTGATGAATCAGTTGGAAGAGGATCTTGCTCAGCAAGAGA 1740  
QY 1759 CGGAGTGAATCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1818  
DB 1741 CGGAGTGAATCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1800  
QY 1819 CGGAAGCGCAGAAATCTGACCAATAAATCTTGAAGCTTAAGGATCAAGGGAAGCTGAA 1878  
DB 1801 CGGAAGCGCAGAAATCTGACCAATAAATCTTGAAGCTTAAGGATCAAGGGAAGCTGAA 1860  
QY 1879 GTGGGAATATGCAAACTGGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938  
DB 1861 GTGGGAATATGCAAACTGGAGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
QY 1939 CTCGAAGAGAACTGGAGAAGGCT----- 1962  
DB 1921 CTCGAAGAGAACTGGAGAAGGCTGTAAGAGCCAGCACGAGGCCCAGAGCTGCTGCAG 1980  
QY 1963 -----GCAAGAGCGAGCCGAGAGGAGCTGAGAGAGCTGAGAAACCCGAGAG 2010  
DB 1981 AATATCCGCGAGGCAAGAGGCGAGCCGAGAGGAGCTGAGAGAGCTGCAAGAACCCGAGAG 2040  
QY 2011 GATTCTTCTGAAGGCAATCAGAAAGAGCTGCTGGAAGCTGAGGAACCGCCGCTTCTG 2070  
DB 2041 GATTCTTCTGAAGGCAATCAGAAAGAGCTGCTGGAAGCTGAGGAACCGCCGCTTCTG 2100  
QY 2071 GAGAACAGGTAAGAGACTAGAGACCATGAGCGCTAGAGAAACAGACTGAAGGATGAC 2130  
DB 2101 GAGAACAGGTAAGAGACTAGAGACCATGAGCGCTAGAGAAACAGACTGAAGGATGAC 2160  
QY 2131 ATCCAGACAAATATCCCAACAGATCCAGCAGATGGCTGATAAAATCTGAGAGCTCGAAGAG 2190  
DB 2161 ATCCAGACAAATATCCCAACAGATCCAGCAGATGGCTGATAAAATCTGAGAGCTCGAAGAG 2220  
QY 2191 AAACATCGGAGGCCCCAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG 2250  
DB 2221 AAACATCGGAGGCCCCAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG 2280  
QY 2251 CAGCACTATCAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2310  
DB 2281 CAGCACTATCAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2340  
QY 2311 AAGGAGACACTGAGAAACATGATGAGAGACAGAGGAGGAGGCCCATGAGAAAGGCAAA 2370  
DB 2341 AAGGAGACACTGAGAAACATGATGAGAGACAGAGGAGGAGGCCCATGAGAAAGGCAAA 2400  
QY 2371 ATTCTCAGCGAAACAGAAAGGAGTGAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2430  
DB 2401 ATTCTCAGCGAAACAGAAAGGAGTGAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2460  
QY 2431 CAGAGGATGTGAACTGTCTGAAGCCAAATAAATTGCGAGCAAAATAGCAGTCTTTTACC 2490  
DB 2461 CAGAGGATGTGAACTGTCTGAAGCCAAATAAATTGCGAGCAAAATAGCAGTCTTTTACC 2520  
QY 2491 CRAAGGAACATGAGGCCCAAGAGAGATGATTTCTGAATCTCAGCAACAGAAATTTTAC 2550  
DB 2521 CALAGGAACATGAGGCCCAAGAGAGATGATTTCTGAATCTCAGCAACAGAAATTTTAC 2580  
QY 2551 CTGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCCGAAACTTGGAGAGCAGCTGGAG 2610  
DB 2581 CTGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCCGAAACTTGGAGAGCAGCTGGAG 2640  
QY 2611 AAGATCAGCCACCAAGACCAACAGTGAAGATCGGCTCTGGAATCTGGAACAGAGATG 2670  
DB 2641 AAGATCAGCCACCAAGACCAACAGTGAAGATCGGCTCTGGAATCTGGAACAGAGATG 2700  
QY 2671 CGGAGGTCAGTCTAGAGCACAGAGAGCAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2730  
DB 2701 CGGAGGTCAGTCTAGAGCACAGAGAGCAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2760  
QY 2731 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGACGGGCG 2790

DB 2761 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGACAGCCCTGACGCTGACGGGCG 2820  
QY 2791 GCCCTGGAGAGCCAGCTTCGCGAGCGAAGACAGAGCTTGGAGAGACCCACAGCAGAAGCT 2850  
DB 2821 GCCCTGGAGAGCCAGCTTCGCGAGCGAAGACAGAGCTTGGAGAGACCCACAGCAGAAGCT 2880  
QY 2851 GAAGAGAGATCCAGGCACTCAGCGCATAGAGATGAATCCAGCGCAAAATTTGATGCT 2910  
DB 2881 GAAGAGAGATCCAGGCACTCAGCGCATAGAGATGAATCCAGCGCAAAATTTGATGCT 2940  
QY 2911 CTTCTTAACAGCTGTACTGTAAATCAGACCTGAGGAGCAGCTTAAACCCAGCTGACCGAG 2970  
DB 2941 CTTCTTAACAGCTGTACTGTAAATCAGACCTGAGGAGCAGCTTAAACCCAGCTGACCGAG 3000  
QY 2971 GACAAAGCTGAATCAACCAACAAACTTTCTAATTGCTCCAAACAACTCGATGAGGCTTCT 3030  
DB 3001 GACAAAGCTGAATCAACCAACAAACTTTCTAATTGCTCCAAACAACTCGATGAGGCTTCT 3060  
QY 3031 GGCGCAACGACGAGATTGTACAACTGGAAGTGAAGTGAACCATCTCCGCGGAGATC 3090  
DB 3061 GGCGCAACGACGAGATTGTACAACTGGAAGTGAAGTGAACCATCTCCGCGGAGATC 3120  
QY 3091 ACGGAAACGAGAGATGCAAGCTTACGAGCCAGAAACGATGGAGGCTCTGAAAGCACG 3150  
DB 3121 ACGGAAACGAGAGATGCAAGCTTACGAGCCAGAAACGATGGAGGCTCTGAAAGCACG 3180  
QY 3151 TGCACCACTGCTGGAGGAACAGGTCATGGAATGAGGCGCTTAACGATGAGCTGTAGAA 3210  
DB 3181 TGCACCACTGCTGGAGGAACAGGTCATGGAATGAGGCGCTTAACGATGAGCTGTAGAA 3240  
QY 3211 AAAGAGCGGAGTGGAGGCGCTGAGGAGGCTCTGAGGAGGCTGATGAGAAATCCGATTTGAG 3270  
DB 3241 AAAGAGCGGAGTGGAGGCGCTGAGGAGGCTCTGAGGAGGCTGATGAGAAATCCGATTTGAG 3300  
QY 3271 TGTGGGTTGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCAGGCGGAGGCC 3330  
DB 3301 TGTGGGTTGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCAGGCGGAGGCC 3360  
QY 3331 GATCAGCGGATCACCGAGTCTCGCAGGCTGGAGGCTGGAGTGGAGGAGCAAGGCT 3390  
DB 3361 GATCAGCGGATCACCGAGTCTCGCAGGCTGGAGGCTGGAGTGGAGGAGCAAGGCT 3420  
QY 3391 GAGATTCTGCTGACAGGCTCTCAAGAGCAGAGCTGAAGGCGGAGGAGCTCTCT 3450  
DB 3421 GAGATTCTGCTGACAGGCTCTCAAGAGCAGAGCTCTCAAGAGCAGAGGCTCTCT 3480  
QY 3451 GACAAGCTCAATGACCTGGAGAGAGAGCTGCTGCTGAAATGAATGCCGAGCTTA 3510  
DB 3481 GACAAGCTCAATGACCTGGAGAGAGAGCTGCTGCTGAAATGAATGCCGAGCTTA 3540  
QY 3511 CAGCAGAAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3570  
DB 3541 CAGCAGAAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3600  
QY 3571 TTACAGCAGCAGATGGACCTGCAAGAAATCAATTTCCGCTGCTCAGGAGCTGCAA 3630  
DB 3601 TTACAGCAGCAGATGGACCTGCAAGAAATCAATTTCCGCTGCTCAGGAGCTGCAA 3660  
QY 3631 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGTGAATTTGGAGTATCAGCTG 3690  
DB 3661 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGTGAATTTGGAGTATCAGCTG 3720  
QY 3691 GAAACATTCAGTCTCTATTCTCTATGAAAGGTAAGTGAAGGCACTATTTCTCAA 3750  
DB 3721 GAAACATTCAGTCTCTATTCTCTATGAAAGGTAAGTGAAGGCACTATTTCTCAA 3780  
QY 3751 CAAACCAAACTCAATGATTTCTGCAAGCCAAATGGACCACTCTAAAGAAAGAAAG 3810  
DB 3781 CAAACCAAACTCAATGATTTCTGCAAGCCAAATGGACCACTCTAAAGAAAGAAAG -- 3838  
QY 3811 GGTTTATTAGTCGAGGAAAGAGGACCTGCTTTTACCACACAGAGTTCTCTGCGATGAC 3870

Db 3839 -----AGTTTCTCTGCAGTAC 3855  
QY 3871 AATGAGCTGAAGTGGCCCTTGAGAAAGAGAAAGCTCGCTGTGCAGAGCTAGAGAGGCC 3930  
Db 3856 AATGAGCTGAAGTGGCCCTTGAGAAAGAGAAAGCTCGCTGTGCAGAGCTAGAGAGGCC 3915  
QY 3931 CTTTCAAGAGACCCGATCGAGCTCCGGTCCGGCCCGGAGGAAGCTGCCCAACGCAAAAGCA 3990  
Db 3916 CTTTCAAGAGACCCGATCGAGCTCCGGTCCGGCCCGGAGGAAGCTGCCCAACGCAAAAGCA 3975  
QY 3991 ACGGACCAACCCACACCCATPCCAGCCAGCCACCGCGAGGAGCAGATCGCATGTCCGCC 4050  
Db 3976 ACGGACCAACCCACACCCATPCCAGCCAGCCACCGCGAGGAGCAGATCGCATGTCCGCC 4035  
QY 4051 ATCGTGGCGTCCGAGAGCAGCCAGCCAGTGCATGAGCTCTCGCCCGCCGCAATCCAGC 4110  
Db 4036 ATCGTGGCGTCCGAGAGCAGCCAGCCAGTGCATGAGCTCTCGCCCGCCGCAATCCAGC 4095  
QY 4111 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCCGGCGTCTTAAAGAAAGCATGCAC 4170  
Db 4096 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCCGGCGTCTTAAAGAAAGCATGCAC 4155  
QY 4171 CACAATATTTCTTCAACCGATCAACGTAGAGCTGAACATGCGAGCCACAAAGTGTCTGTG 4230  
Db 4156 CACAATATTTCTTCAACCGATCAACGTAGAGCTGAACATGCGAGCCACAAAGTGTCTGTG 4215  
QY 4231 TGTCTGGATACCGTGCACCTTGGAGCCAGGCATCCAAATGTCTCGAATGTTCAGGTGATG 4290  
Db 4216 TGTCTGGATACCGTGCACCTTGGAGCCAGGCATCCAAATGTCTCGAATGTTCAGGTGATG 4275  
QY 4291 TGTACCCCAAGTGTCTCAAGTGTTCGAGCCACCTGGGGTTCGCTGCAATATGCC 4350  
Db 4276 TGTACCCCAAGTGTCTCAAGTGTTCGAGCCACCTGGGGTTCGCTGCAATATGCC 4335  
QY 4351 ACACACTTCCACGAGGCCCTTTCGCGTGACAAATGAACTCCCGAGGCTCCAGACCAAG 4410  
Db 4336 ACACACTTCCACGAGGCCCTTTCGCGTGACAAATGAACTCCCGAGGCTCCAGACCAAG 4395  
QY 4411 GAGCCACAGCAGCTTGACCTTGGAGGGTGGATGAAGGTCCCGAGGAATTAACAAAGCA 4470  
Db 4396 GAGCCACAGCAGCTTGACCTTGGAGGGTGGATGAAGGTCCCGAGGAATTAACAAAGCA 4455  
QY 4471 GGACAGAGGCTGGACAGGAGTACATTTGCTGGAGGATCAAAAGTCTCATTTAT 4530  
Db 4456 GGACAGAGGCTGGACAGGAGTACATTTGCTGGAGGATCAAAAGTCTCATTTAT 4515  
QY 4531 GACATGAAGCCAGAGAGCTGACAGAGCCGGTGGAAAGATTTGAGCTGTCCTTCC 4590  
Db 4516 GACATGAAGCCAGAGAGCTGACAGAGCCGGTGGAAAGATTTGAGCTGTCCTTCC 4575  
QY 4591 GACGGGATGTATCTATTATGTCGCTGGTGGTTCGGAATTCGCAATACAGCCAA 4650  
Db 4576 GACGGGATGTATCTATTATGTCGCTGGTGGTTCGGAATTCGCAATACAGCCAA 4635  
QY 4651 GCA----- 4653  
Db 4636 GCAGATGTCCCATACATCTGAAGATGGAAATCTCACCCGCACACACCTGCTGGCCGGG 4695  
QY 4654 ----- 4653  
Db 4696 AGAACCCTTACTTGTAGCTCCAGCTTCCCTTGACAAAACAGCGCTGGGTACCCGCTTA 4755  
QY 4654 -----GAAAGAGAGAGCTGATGCTAAACTG 4680  
Db 4756 GAATCAGTTGTCGAGTGGGAGAGTTTCTAGGGGAAAGAGAGAGTGTATGCTAAACTG 4815  
QY 4681 CTTTGGAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAATGCAAGCTG 4740  
Db 4816 CTTTGGAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAATGCAAGCTG 4875  
QY 4741 CCCTTCAGTGCACAGGTGTGTGTGGGACCCAGAGGAGGCTCTACGCGCTGAATGTC 4800  
Db 4876 CCCTTCAGTGCACAGGTGTGTGTGGGACCCAGAGGAGGCTCTACGCGCTGAATGTC 4935

QY 4801 TTGAAAACTCCCTAAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAAATTTATATTATC 4860  
Db 4936 TTGAAAACTCCCTAAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAAATTTATATTATC 4995  
QY 4861 AAGGACCTGAGAGAGCTATCTATGATAGAGAGAGAGCGGCACTGTGTCTTTGGGAC 4920  
Db 4996 AAGGACCTGAGAGAGCTATCTATGATAGAGAGAGAGCGGCACTGTGTCTTTGGGAC 5055  
QY 4921 GTGAAGAAAGTGAAGACAGTCCCTGGCCAGTCCCACTGCTGCTCCAGCCGACATCTCA 4980  
Db 5056 GTGAAGAAAGTGAAGACAGTCCCTGGCCAGTCCCACTGCTGCTCCAGCCGACATCTCA 5115  
QY 4981 CCCAAACATTTTGAAGGTGTCAAGGGGTGCCACTTTGTTGGGGAGGCAAGATTGAGAAC 5040  
Db 5116 CCCAAACATTTTGAAGGTGTCAAGGGGTGCCACTTTGTTGGGGAGGCAAGATTGAGAAC 5175  
QY 5041 GGGCTCTGCAATCTGTGAGCCATGCCAGCAAAAGTGTGTCATTTCTCGCTACAAACGAAAC 5100  
Db 5176 GGGCTCTGCAATCTGTGAGCCATGCCAGCAAAAGTGTGTCATTTCTCGCTACAAACGAAAC 5235  
QY 5101 CTCAGCAAAATAGTGCATPCCGAAAGAGATAGACACTTCAGAGCCCTGCAGCTGTATCCAC 5160  
Db 5236 CTCAGCAAAATAGTGCATPCCGAAAGAGATAGACACTTCAGAGCCCTGCAGCTGTATCCAC 5295  
QY 5161 TTCAACAAATACAGTATCTCTATTTGAAACCAATAAATTTACGAAATCGACATGAAGCAG 5220  
Db 5296 TTCAACAAATACAGTATCTCTATTTGAAACCAATAAATTTACGAAATCGACATGAAGCAG 5355  
QY 5221 TACAGCTCGAGGAATTTCTTGGATAAGATGACCAATTCCTTGGCACCTGTGTCTTTGCC 5280  
Db 5356 TACAGCTCGAGGAATTTCTTGGATAAGATGACCAATTCCTTGGCACCTGTGTCTTTGCC 5415  
QY 5281 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAAACAGGCGAGGCGACGAGAG 5340  
Db 5416 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAAACAGGCGAGGCGACGAGAG 5475  
QY 5341 GAGTACTTGTGTCTTCCAGATTTGGAGTGTCTGTGATTTCTACGGAAGAGCTAGC 5400  
Db 5476 GAGTACTTGTGTCTTCCAGATTTGGAGTGTCTGTGATTTCTAGGAAGAGCTAGC 5535  
QY 5401 CGCACAGACGATCTCAAGTGGAGTTCCTTACCTTTTGGCCTTACAGAGAACCCCTAT 5460  
Db 5536 CGCACAGACGATCTCAAGTGGAGTTCCTTACCTTTTGGCCTTACAGAGAACCCCTAT 5595  
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTCAGATCCAGGCGACGCTCCTCAGCA 5520  
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTCAGATCCAGGCGACGCTCCTCAGCA 5655  
QY 5521 GGGACCCCTGCGCGAGCGTACCTGGAATCCCGAAACCCGGCTACCTGGGCGCTGCCATT 5580  
Db 5656 GGGACCCCTGCGCGAGCGTACCTGGAATCCCGAAACCCGGCTACCTGGGCGCTGCCATT 5715  
QY 5581 TCCTCAGAGGAGTATTTACTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGTGTC 5640  
Db 5716 TCCTCAGAGGAGTATTTACTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGTGTC 5775  
QY 5641 AAGGAAACCTTGTGAAGGAGTCCGGCACTGAACACACCGGGGCGCTCCACCTCCCGC 5700  
Db 5776 AAGGAAACCTTGTGAAGGAGTCCGGCACTGAACACACCGGGGCGCTCCACCTCCCGC 5835  
QY 5701 AGCAGCCCCAAACAGGAGGCCACCCCACTGATCAACAGAGCAATCAACAGCGGTGCC 5760  
Db 5836 AGCAGCCCCAAACAGGAGGCCACCCCACTGATCAACAGAGCAATCAACAGCGGTGCC 5895  
QY 5761 TCCAGCCAGCGCCCGGAGGCCCGCAGCCACCGGAGGCGCAAGCACCACCCACCGC 5820  
Db 5896 TCCAGCCAGCGCCCGGAGGCCCGCAGCCACCGGAGGCGCAAGCACCACCCACCGC 5955  
QY 5821 TACCGCGAGGGCGGACCGAGCTGCGAGGAGCAAGTCTCTGTGGCCGCCCTTGGAGCGA 5880  
Db 5956 TACCGCGAGGGCGGACCGAGCTGCGAGGAGCAAGTCTCTGTGGCCGCCCTTGGAGCGA 6015

QY 5981 GAGAAAGTCCCGCCGCGGAGTCTCAGCACCGCGGAGAGAGCGGTCCCGCGGAGAGCGTGT 5940  
 Db 6016 GAGAAGTCCCGCCGCGGAGTCTCAGCACCGCGGAGAGAGCGGTCCCGCGGAGAGCGTGT 6075  
 QY 5941 GAAGACAGCAGCAGCGGCGCGGCTCCCTGCGGAGCGGTGAGGACCCCGCTGTCCCGAGGTG 6000  
 Db 6076 GAAGACAGCAGCAGCGGCGCGGCTCCCTGCGGAGCGGTGAGGACCCCGCTGTCCCGAGGTG 6135  
 QY 6001 AACAAAGGAAGAGCGCAGAGTGC 6023  
 Db 6136 AACAAAGTCTGGACCAAGTCTTC 6158

RESULT 7  
 AAS06701  
 ID AAS06701 standard; cDNA; 6159 BP.  
 XX AC  
 XX AC  
 XX AC  
 DT 12-SEP-2001 (first entry)  
 XX DE  
 XX DE  
 XX DE  
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
 KW metabolic disorder; immune related disease; neurological disorder;  
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
 KW reproductive disorder; gene therapy; ss.  
 XX OS  
 XX Homo sapiens.  
 XX PN  
 XX W0200138503-A2.  
 XX PD  
 XX 31-MAY-2001.  
 XX PF  
 XX 22-NOV-2000; 2000WO-US032085.  
 XX PR  
 XX 24-NOV-1999; 99US-0167482P.  
 XX PA  
 XX (SUGEN-) SUGEN INC.  
 XX PI  
 XX Plowman GD, Whyte D, Manning G, Sudareanam S, Martinez R;  
 XX PI Flanagan P, Clary D;  
 XX XX  
 DR WPI; 2001-343950/36.  
 DR P-PSDB; AAU03501.  
 XX  
 XX Nucleic acids encoding human kinase polypeptides, useful for preventing  
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
 PT neuronal-associated diseases, and microbial infections.  
 XX  
 PS Example 1; Fig 1; 433pp; English.

CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel  
 CC protein kinases have been identified as members of the tyrosine or  
 CC serine/threonine kinase (PTK and STK) families. The polynucleotides  
 CC encoding protein kinases and the polypeptides may be used in the  
 CC prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate kinase expression. For example, they may be used to treat  
 CC cancers (especially cancers of haematopoietic origin), cardiovascular  
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
 CC immune related diseases (e.g. rheumatoid arthritis), neurological  
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
 CC Additionally, polynucleotides encoding protein kinases may be used for  
 CC gene therapy and as DNA probes in diagnostic assays. The protein kinase  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC against the protein kinases and in assays to identify modulators of  
 CC protein kinase expression and activity

SQ Sequence 6159 BP; 1732 A; 1549 C; 1680 G; 1198 T; 0 U; 0 Other;  
 Query Match 86.0%; Score 5650.4; DB 4; Length 6159;

Best Local Similarity 95.9%; Pred. No. 0;  
 Matches 5943; Conservative 0; Mismatches 11; Indels 246; Gaps 4;

QY 19 ATGTTGAAGTTCAAATATATGGAGCGGGAATCCTTTGGATGCTGGTGTCTGAACCCATT 78  
 Db 1 ATGTTGAAGTTCAAATATATGGAGCGGGAATCCTTTGGATGCTGGTGTCTGAACCCATT 60  
 QY 79 GCCAGCCGGGCTCCAGGCTGAATCTGTTTCCAGGGGAAACCACTTTTATGACTCAA 138  
 Db 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTTCCAGGGGAAACCACTTTTATGACTCAA 120  
 QY 139 CAGCAGATGTCCTCTTTTCCGAGAGAGGATATTAGATGCCCTCTTGTCTCTTTGAA 198  
 Db 121 CAGCAGATGTCCTCTCTTTTCCGAGAGAGGATATTAGATGCCCTCTTGTCTCTTTGAA 180  
 QY 199 GAATGCAGTCAGCCTGCTCTGATGAAGATTAAAGCACCTGAGCAACTTTGTCCGGAAGTG 255  
 Db 181 GAATGCAGTCAGCCTGCTCTGATGAAGATTAAAGCACCTGAGCAACTTTGTCCGGAAGTG 240  
 QY 256 TATTCGACACCATAGCTGAGTTACAGAGCTCAGCCTTCGGCAAAAGAGACTTCGAAGTC 315  
 Db 241 TATTCGACACCATAGCTGAGTTACAGAGCTCAGCCTTCGGCAAAAGAGACTTCGAAGTC 300  
 QY 316 AGAAGTCTTGTAGTGTGTGTCCTTTGCTGAAGTGCAGTGTGAAGAGAAAGCAACC 375  
 Db 301 AGAAGTCTTGTAGTGTGTGTCCTTTGCTGAAGTGCAGTGTGAAGAGAAAGCAACC 360  
 QY 376 GGGACATCTATGCTATGAAAGTGATGAAGAAGGCTTTATTTGGCCAGGAGCAGGTT 435  
 Db 361 GGGACATCTATGCTATGAAAGTGATGAAGAAGGCTTTATTTGGCCAGGAGCAGGTT 420  
 QY 436 TCATTTTGGAGGAGAGCGGAACATATTATCTCGAAGACAAAGCCGTCGATCCGCCAA 495  
 Db 421 TCATTTTGGAGGAGAGCGGAACATATTATCTCGAAGACAAAGCCGTCGATCCGCCAA 480  
 QY 496 TTACAGTATGCTTTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGCCTGGA 555  
 Db 481 TTACAGTATGCTTTTCAGGACAAAATACCTTTATCTGATGGGAATATCAGCCTGGA 540  
 QY 556 GGGACATGCTGTCACCTTTTGAATAGATATGAGACAGTGTAGATGAAGAACTGATACAG 615  
 Db 541 GGGACATGCTGTCACCTTTTGAATAGATATGAGACAGTGTAGATGAAGAACTGATACAG 600  
 QY 616 TTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGAGCTTCATCTGATGGGATACGTGCAT 675  
 Db 601 TTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGAGCTTCATCTGATGGGATACGTGCAT 660  
 QY 676 CGAGACATCAAGCTGAGACATCTCGTTGACCGACAGGACACATCAAGCTGGTGGAT 735  
 Db 661 CGAGACATCAAGCTGAGACATCTCGTTGACCGACAGGACACATCAAGCTGGTGGAT 720  
 QY 736 TTTTGATCTCCGCGAAATGAATTCAAACAAGATGTTGAATGCCAACTCCCGATTGGG 795  
 Db 721 TTTTGATCTCCGCGAAATGAATTCAAACAAGATGTTGAATGCCAACTCCCGATTGGG 780  
 QY 796 ACCCAGATTACATGGCTCTCGAAGTGTGACTGTGATGAACGGGATGGAAGGACCC 855  
 Db 781 ACCCAGATTACATGGCTCTCGAAGTGTGACTGTGATGAACGGGATGGAAGGACCC 840  
 QY 856 TACGGCTGAGCTGTGACTGCTGCTGAGTGGGCTGATTTGCCCTATCAGATGATTTATGGG 915  
 Db 841 TACGGCTGAGCTGTGACTGCTGCTGAGTGGGCTGATTTGCCCTATCAGATGATTTATGGG 900  
 QY 916 AGATGCCCTTCGACAGGGAACCTCTGCGAAGAACCTTCAATAACATTATGAATTTCCAG 975  
 Db 901 AGATGCCCTTCGACAGGGAACCTCTGCGAAGAACCTTCAATAACATTATGAATTTCCAG 960  
 QY 976 CGGTTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTCTTCTTCTGATCTGATCAA 1035  
 Db 961 CGGTTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTCTTCTTCTGATCTGATCAA 1020  
 QY 1036 AGCTGTGTGCGGCCAGAAAGAGACTGAAAGTTTGAAGGCTTTTGTGCGCATCTTTC 1095

Db 1021 AGCTTGTTGTGGGCCAGNAGAGAGACTGAAGTTTGAAGGTCCTTGTGTCGCATCCTTTC 1080  
QY 1096 TTCTCTAAAAATTGACTGGAAACAACTTCGTAACTCTCTCCCTCCCTTCCTGTTCCACCCCTC 1155  
Db 1081 TTCTCTAAAAATTGACTGGAAACAACTTCGTAACTCTCTCCCTCCCTTCCTGTTCCACCCCTC 1140  
QY 1156 AAGTCTGACGATGACACTTCCAAATTTTGTGAACACAGAGAGAAATTCGTGGGTTTCATCC 1215  
Db 1141 AAGTCTGACGATGACACTTCCAAATTTTGTGAACACAGAGAGAAATTCGTGGGTTTCATCC 1200  
QY 1216 TCTCCGTGCCAGTCCAGCCCTCAGGCTTCTCGGTGAAGAACTCCGCTTTGTGGGTTT 1275  
Db 1201 TCTCCGTGCCAGTCCAGCCCTCAGGCTTCTCGGTGAAGAACTCCGCTTTGTGGGTTT 1260  
QY 1276 TCGTACAGCAAGCACTGGGATTCCTTGGTAGATCTGAGTCTGTTGTCTCGGCTCGAC 1335  
Db 1261 TCGTACAGCAAGCACTGGGATTCCTTGGTAGATCTGAGTCTGTTGTCTCGGCTCGAC 1320  
QY 1336 TCCCTTGCCAAAGACTAGTCCATGGAAAGAACTTCTCATCAAAAGCAAGAGACTACAA 1395  
Db 1321 TCCCTTGCCAAAGACTAGTCCATGGAAAGAACTTCTCATCAAAAGCAAGAGACTACAA 1380  
QY 1396 GACTCTCAGGACAAAGTGTCAAGATGGAGCAGGAAATGACCCGTTTACATCGGAGATG 1455  
Db 1381 GACTCTCAGGACAAAGTGTCAAGATGGAGCAGGAAATGACCCGTTTACATCGGAGATG 1440  
QY 1456 TCAGAGTGGAGGCTGTGCTTAGTCAAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAG 1515  
Db 1441 TCAGAGTGGAGGCTGTGCTTAGTCAAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAG 1500  
QY 1516 AGATCCCTCTCGAGCAGGACCTTGCTACTCATCATCAGAAATGCAAGTGTGAGTTAAAGCGA 1575  
Db 1501 AGATCCCTCTCGAGCAGGACCTTGCTACTCATCATCAGAAATGCAAGTGTGAGTTAAAGCGA 1560  
QY 1576 AGTTTGGACCAAGCAGGATGAGTGTCCAGGAGGATGACAAAGCACTGCAGCTTCTC 1635  
Db 1561 AGTTTGGACCAAGCAGGATGAGTGTCCAGGAGGATGACAAAGCACTGCAGCTTCTC 1620  
QY 1636 CATGATATCAGAGAGCAGACCGGAGCTCCAGAAATCAAAAGCAGGAGTACAGGCT 1695  
Db 1621 CATGATATCAGAGAGCAGACCGGAGCTCCAGAAATCAAAAGCAGGAGTACAGGCT 1680  
QY 1696 CAAAGTGGAAAGAAATGAGTTGATGATGATCAATCAGTTGGAAGAGATCTTGTCTCACAAGA 1755  
Db 1681 CAAAGTGGAAAGAAATGAGTTGATGATGATCAATCAGTTGGAAGAGATCTTGTCTCACAAGA 1740  
QY 1756 AGACGGAGTGATCTCTACCAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATTC 1815  
Db 1741 AGACGGAGTGATCTCTACCAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATTC 1800  
QY 1816 AAGCGGAAAGCGACAGAAATGTGAGCATAAACTGTTGAAGGCTAAGATCAAGGGAAGCT 1875  
Db 1801 AAGCGGAAAGCGACAGAAATGTGAGCATAAACTGTTGAAGGCTAAGATCAAGGGAAGCT 1860  
QY 1876 GAAGTGGGAGATATCGGAACTGGAGAGATCAATGCTGACGACAGCTCAAAATTCAG 1935  
Db 1861 GAAGTGGGAGATATCGGAACTGGAGAGATCAATGCTGACGACAGCTCAAAATTCAG 1920  
QY 1936 GAGCTCCAAAGAGAACTGGAGAGGCT----- 1962  
Db 1921 GAGCTCCAAAGAGAACTGGAGAGGCTGTAAAGCCAGCACGAGGCCACCGAGTGCTG 1980  
QY 1963 -----GCAAGGAGCGAGCGGAGAGGAGCTGGAGAGCTGGAGACCGTGCAGAACCGA 2007  
Db 1981 CAGAAATATCCGCGAGGCAAGGAGCGGAGGAGCTGGAGAGCTGGAGAGCTGCAGAACCGA 2040  
QY 2008 GAGGATCTTCTGAAGGCAATCAGAAAGAGCTGGTGAAGCTGAGGAAACCGCCCATCTC 2067  
Db 2041 GAGGATCTTCTGAAGGCAATCAGAAAGAGCTGGTGAAGCTGAGGAAACCGCCCATCTC 2100  
QY 2068 CTGGAGAACAAAGTAAAGAGACTAGAGACCATGAGCGGTAGAGAAACAGACTGAGGAT 2127  
Db 2101 CTGGAGAACAAAGTAAAGAGACTAGAGACCATGAGCGGTAGAGAAACAGACTGAGGAT 2160

QY 2128 GACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAAATTCGGAGCTCGAA 2187  
Db 2161 GACATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAAATTCGGAGCTCGAA 2220  
QY 2188 GAGAAACATCGGAGGCCCAAAGTCTCAGCCACGACCTAGAAGTGCACTGAAACAGAAA 2247  
Db 2221 GAGAAACATCGGAGGCCCAAAGTCTCAGCCACGACCTAGAAGTGCACTGAAACAGAAA 2280  
QY 2248 GAGCAGCACTATGAGGAAAGATTAAGTCTTGGCAATCAGATAAAGAAAGACCTGGCT 2307  
Db 2281 GAGCAGCACTATGAGGAAAGATTAAGTCTTGGCAATCAGATAAAGAAAGACCTGGCT 2340  
QY 2308 GACAAGGAGACACTGGGAAACATGATCGAGACACAGAGGAGGCCCCATGAGAAGGC 2367  
Db 2341 GACAAGGAGACACTGGGAAACATGATCGAGACACAGAGGAGGCCCCATGAGAAGGC 2400  
QY 2368 AAAATTCAGCGAAACAGAAAGGCGATGATCAATGCTATGATTCGAAGATCAGATCCCTG 2427  
Db 2401 AAAATTCAGCGAAACAGAAAGGCGATGATCAATGCTATGATTCGAAGATCAGATCCCTG 2460  
QY 2428 GAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTCAGCAAAATAGCAGTCTTTTT 2487  
Db 2461 GAACAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTCAGCAAAATAGCAGTCTTTTT 2520  
QY 2488 ACCCAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTT 2547  
Db 2521 ACCCAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTT 2580  
QY 2548 TACCTGAGACACAGGCTGGGAAAGTTGGAGGCCACAGAACCCGAAACTGGAGGAGCAGCTG 2607  
Db 2581 TACCTGAGACACAGGCTGGGAAAGTTGGAGGCCACAGAACCCGAAACTGGAGGAGCAGCTG 2640  
QY 2608 GAGAAGATCAGCCACCAAGACCACAGTGACAAGAAATCGGCTCTGGAATCTGGAGACAAGA 2667  
Db 2641 GAGAAGATCAGCCACCAAGACCACAGTGACAAGAAATCGGCTCTGGAATCTGGAGACAAGA 2700  
QY 2668 TTGCGGAGGTGAGTCTAGAGCAGGAGGAGCAAACTGAGGCTCAGCGCCAGCTCACA 2727  
Db 2701 TTGCGGAGGTGAGTCTAGAGCAGGAGGAGCAAACTGAGGCTCAGCGCCAGCTCACA 2760  
QY 2728 GAGCTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCCCTGCAAGGCTGCACGG 2787  
Db 2761 GAGCTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCCCTGCAAGGCTGCACGG 2820  
QY 2788 GCGGCCCTGAGAGCCAGCTTCGCGCAAGACAGAGCTGGAAGAGACACAGAGCAAA 2847  
Db 2821 GCGGCCCTGAGAGCCAGCTTCGCGCAAGACAGAGCTGGAAGAGACACAGAGCAAA 2880  
QY 2848 GCTGAAGAGGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAAAATTTGAT 2907  
Db 2881 GCTGAAGAGGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAAAATTTGAT 2940  
QY 2908 GCTCTTCGTAAACAGCTGACTGTAAATCAGAGACCTGGAGGAGCAGCTAAACCCAGCTGACC 2967  
Db 2941 GCTCTTCGTAAACAGCTGACTGTAAATCAGAGACCTGGAGGAGCAGCTAAACCCAGCTGACC 3000  
QY 2968 GAGGACAAACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGCT 3027  
Db 3001 GAGGACAAACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGCT 3060  
QY 3028 TCTGGCCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAAGCTATCCCGCCGGAG 3087  
Db 3061 TCTGGCCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAAGCTATCCCGCCGGAG 3120  
QY 3088 ATCAGGAAACGAGAGATGCGCTTACAGCCAGAGCAAAAGATGAGGCTCTGAAGACC 3147  
Db 3121 ATCAGGAAACGAGAGATGCGCTTACAGCCAGAGCAAAAGATGAGGCTCTGAAGACC 3180  
QY 3148 ACCTGCAACCTGCTGGAGGACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGCTA 3207  
Db 3181 ACCTGCAACCTGCTGGAGGACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGCTA 3240

QY 3208 GAAAAAGCGGCGAGTGGGAGGCTGAGAGAGCGTCTCGGTCATGAGAAATCCCAAGTTT 3267  
Db 3241 GAAAAAGCGGCGAGTGGGAGGCTGAGAGAGCGTCTCGGTCATGAGAAATCCCAAGTTT 3300  
QY 3268 GAGTGTCCGGTTCGAGAGCTGAGAGAAATGTCGACACCCAGAGAAACAGACAGCGCGAGA 3327  
Db 3301 GAGTGTCCGGTTCGAGAGCTGAGAGAAATGTCGACACCCAGAGAAACAGACAGCGCGAGA 3360  
QY 3328 GCCATCAGCGGATCAACCGAGTCTCGCCAGGTGTGAGCTGCGAGTGAAGAGAGCAAG 3387  
Db 3361 GCCATCAGCGGATCAACCGAGTCTCGCCAGGTGTGAGCTGCGAGTGAAGAGAGCAAG 3420  
QY 3388 GCTCAGATCTCGCTCTGACAGAGCTCTCAAGAGCAGAGCTGAAGCCCGAGAGCCCTC 3447  
Db 3421 GCTCAGATCTCGCTCTGACAGAGCTCTCAAGAGCAGAGCTGAAGCCCGAGAGCCCTC 3480  
QY 3448 TCTGACAGCTCAATGACCTGGAGAGAGCTGATGCTTGAATGAATGATGCCGAAGC 3507  
Db 3481 TCTGACAGCTCAATGACCTGGAGAGAGCTGATGCTTGAATGAATGATGCCGAAGC 3540  
QY 3508 TTACAGCAGAGCTGGAGCTGACAGAGCTCAACAGAGGCTTCTGAGAGAGAGCC 3567  
Db 3541 TTACAGCAGAGCTGGAGCTGACAGAGCTCAACAGAGGCTTCTGAGAGAGAGCC 3600  
QY 3568 AAATTACAGCAGAGATGAGCTGACGAAATCAATTTTCCGCTGACTCAAGGACTG 3627  
Db 3601 AAATTACAGCAGAGATGAGCTGACGAAATCAATTTTCCGCTGACTCAAGGACTG 3660  
QY 3628 CAAGAGCTTAGATCGGCTGATCTACTGAAGAGAGAGAGTGAAGTGAAGTATCAG 3687  
Db 3661 CAAGAGCTTAGATCGGCTGATCTACTGAAGAGAGAGAGTGAAGTGAAGTATCAG 3720  
QY 3688 CTGGAATCAATTCAGGTTCTTCTATCTCATGAAAGGTGAAATGGAAGCAATTTCT 3747  
Db 3721 CTGGAATCAATTCAGGTTCTTCTATCTCATGAAAGGTGAAATGGAAGCAATTTCT 3780  
QY 3748 CAACAAACCAATCTATGATTTCTGCAAGCAAAATGACCAACCTCTAAAAAGAAA 3807  
Db 3781 CAACAAACCAATCTATGATTTCTGCAAGCAAAATGACCAACCTCTAAAAAGAAA 3840  
QY 3808 AAGGGTTATTTAGTCGACGAGAGAGGACCTGCTTTACACACAGCTTCTCTGAG 3867  
Db 3841 A-----AGTTTCTCTGAG 3855  
QY 3868 TACAATGAGTGAAGTGGCTGGAGAGAGAAAGCTCGTGTGACAGCTAGAGGAA 3927  
Db 3856 TACAATGAGTGAAGTGGCTGGAGAGAGAAAGCTCGTGTGACAGCTAGAGGAA 3915  
QY 3928 GCCCTTCAGAGACCCGATCGAGCTCCGTCGCGCCGGAGGAGAGCTGCCACCGCAA 3987  
Db 3916 GCCCTTCAGAGACCCGATCGAGCTCCGTCGCGCCGGAGGAGAGCTGCCACCGCAA 3975  
QY 3988 GCAACGAGACCAACCAACCCATCCAGCCAGCCACCCGAGGAGCAGATCGCCATGTC 4047  
Db 3976 GCAACGAGACCAACCAACCCATCCAGCCAGCCACCCGAGGAGCAGATCGCCATGTC 4035  
QY 4048 GCCATGCTGGTCGCGAGAGAGCAGAGCCAGTGCATGAGCTGCTGGCCCGCCATCC 4107  
Db 4036 GCCATGCTGGTCGCGAGAGAGCAGAGCCAGTGCATGAGCTGCTGGCCCGCCATCC 4095  
QY 4108 AGCGCAGAGAGAGTCTTCACTCCAGAGGAAATTTAGTCGCGCTTTAAGGAAGCGATG 4167  
Db 4096 AGCGCAGAGAGAGTCTTCACTCCAGAGGAAATTTAGTCGCGCTTTAAGGAAGCGATG 4155  
QY 4168 CACCAAAATTTCTCACCAGTTCAACGTAGAGTGAACATCGAGGAGCAAAAGTGTCT 4227  
Db 4156 CACCAAAATTTCTCACCAGTTCAACGTAGAGTGAACATCGAGGAGCAAAAGTGTCT 4215  
QY 4228 GTGTGTCTGGATACCGTGCATTTTGAGCCAGGATCCAAATGCTCGAATGTCAGGTG 4287  
Db 4216 GTGTGTCTGGATACCGTGCATTTTGAGCCAGGATCCAAATGCTCGAATGTCAGGTG 4275  
QY 4288 ATGTGTACCCCAAGTGTCTCAGCTGTTGCCAGCCACCTGCGGCTGCTGTAATAT 4347

Db 4276 ATGTGTACCCCAAGTGTCTCAGCTGTTGCCAGCCACCTCGGCTTGCCTGCTGAATAT 4335  
QY 4348 GCCACACACTTTCACCGAGGCTTTCGCGCTGACAAAATGAACTCCCAAGTCTCCAGACC 4407  
Db 4336 GCCACACACTTTCACCGAGGCTTTCGCGCTGACAAAATGAACTCCCAAGTCTCCAGACC 4395  
QY 4408 AAGGAGCCACAGCAGAGCTTGCACCTGGAAGGTTGATGAAGTGCACAGGATGAA 4467  
Db 4396 AAGGAGCCACAGCAGAGCTTGCACCTGGAAGGTTGATGAAGTGCACAGGATGAA 4455  
QY 4468 CGAGGACAGCAAGCTGGGACAGGAGTACATTTCTCTGGAGGATCAAAAGTCTCTATT 4527  
Db 4456 CGAGGACAGCAAGCTGGGACAGGAGTACATTTCTCTGGAGGATCAAAAGTCTCTATT 4515  
QY 4528 TATGACAAATGAAGCCAGAGAGCTGGAAGGCGGTGGAAGAAATTTGAGCTGTGCTT 4587  
Db 4516 TATGACAAATGAAGCCAGAGAGCTGGAAGGCGGTGGAAGAAATTTGAGCTGTGCTT 4575  
QY 4588 CCCGAGCGGATGATCTTATTTCTGCTGCTGCTTCCGAACTCGAAATACAGCC 4647  
Db 4576 CCCGAGCGGATGATCTTATTTCTGCTGCTGCTTCCGAACTCGAAATACAGCC 4635  
QY 4648 AAAGCAGA----- 4655  
Db 4636 AAAGCAGATGTCCCATACATACTGAAGTGAATCTCACCGCACACCACTGCTGSCCC 4695  
QY 4656 ----- 4655  
Db 4696 GGGAGAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAGAGCTGGTGAOCGCC 4755  
QY 4656 -----AAAGCAGAGCTGATGCTGCTAAA 4677  
Db 4756 TTAGAAATCAGTTGCGCAGGTGGAGAGTTTCTAGGGAAAGAGCAGAGCTGATGCTAAA 4815  
QY 4678 CTGCTTGGAAATCTCCCTGCTGAAACTGGAAGTGTAGCCGTCTAGACATGAACGCAAG 4737  
Db 4816 CTGCTTGGAAATCTCCCTGCTGAAACTGGAAGTGTAGCCGTCTAGACATGAACGCAAG 4875  
QY 4738 CTGCCCCTTCAAGTACAGGCTGCTGCTGGGACCGAGAGAGGCTCTACGCCCTGAAT 4797  
Db 4876 CTGCCCCTTCAAGTACAGGCTGCTGCTGGGACCGAGAGAGGCTCTACGCCCTGAAT 4935  
QY 4798 GTCTTGAATACTCCCTAACCCATGTCCAGGAATTTGAGCAGTCTTCCAAATTTATAT 4857  
Db 4936 GTCTTGAATACTCCCTAACCCATGTCCAGGAATTTGAGCAGTCTTCCAAATTTATAT 4995  
QY 4858 ATCAAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGCGGACACTGCTCTGTG 4917  
Db 4996 ATCAAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGCGGACACTGCTCTGTG 5055  
QY 4918 GACGTGAAGAAAGTGAACACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATC 4977  
Db 5056 GACGTGAAGAAAGTGAACACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATC 5115  
QY 4978 TCACCCAAATTTTGAAGCTGTCAAGGCTGCACTTGTGTTGGGAGCAGAGTTGAG 5037  
Db 5116 TCACCCAAATTTTGAAGCTGTCAAGGCTGCACTTGTGTTGGGAGCAGAGTTGAG 5175  
QY 5038 AACGGGCTCTGCATCTGTGAGCAGTCCCGCAGAAAGTCTGCTATTCTCCGCTACACGAA 5097  
Db 5176 AACGGGCTCTGCATCTGTGAGCAGTCCCGCAGAAAGTCTGCTATTCTCCGCTACACGAA 5235  
QY 5098 AACCTCAGAAATPACTGATCCCGAAAGAGATGAGACCTCAGAGCCCTGAGCTGTATC 5157  
Db 5236 AACCTCAGAAATPACTGATCCCGAAAGAGATGAGACCTCAGAGCCCTGAGCTGTATC 5295  
QY 5158 CACTTCAACCAATTTACAGTATCTCTTGGAAACCAATTAATTTACGAAATTCAGATGAAG 5217  
Db 5296 CACTTCAACCAATTTACAGTATCTCTTGGAAACCAATTAATTTACGAAATTCAGATGAAG 5355  
QY 5218 CAGTACAGCTCAGGAAATTTCTGGATAAGAAATGACATTTCTGTTGCACTGCTGTGTTT 5277

Db 5356 CAGTACACGCTCGAGGAATTCCTGGATAGAAATGACCATTCCTTGGCACCTCTGTGTTT 5415  
Qy 5278 GCGGCTCTTCAACAGAGCTTCCTCTCAATCGTGCAGGTGAACAGCGCGAGGAGCA 5337  
Db 5416 GCGGCTCTTCAACAGAGCTTCCTCTCAATCGTGCAGGTGAACAGCGCGAGGAGCA 5475  
Qy 5338 GAGGAGTACTTGCTGTGTTTCCACGAATTTGAGTGTTCGTGGATTTCTTACGGAAGACGT 5397  
Db 5476 GAGGAGTACTTGCTGTGTTTCCACGAATTTGAGTGTTCGTGGATTTCTTACGGAAGACGT 5535  
Qy 5398 AGCCGACAGAGATCTCAAGTGGAGTGCCTTTACCTTTGGCCCTTTGCCCTACAGAGAACC 5457  
Db 5536 AGCCGACAGAGATCTCAAGTGGAGTGCCTTTACCTTTGGCCCTTTGCCCTACAGAGAACC 5595  
Qy 5458 TATCTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTTCCTCA 5517  
Db 5596 TATCTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTTCCTCA 5655  
Qy 5518 GAGGAGCCCTTGCCGAGCGTACCTTGAGATCCCGAACCCCGCTACTCTGGGCCCTGCC 5577  
Db 5656 GAGGAGCCCTTGCCGAGCGTACCTTGAGATCCCGAACCCCGCTACTCTGGGCCCTGCC 5715  
Qy 5578 ATTCTCTCAGGAGCGATTTACTTTGGCTCTCATACAGGATAAATTAAGGTCTATTTCG 5637  
Db 5716 ATTCTCTCAGGAGCGATTTACTTTGGCTCTCATACAGGATAAATTAAGGTCTATTTCG 5775  
Qy 5638 TGCAAGGGAACCTCGTGAAGAGTCCGGCACTGAACACCCCGGGCCCGTCCACCTCC 5697  
Db 5776 TGCAAGGGAACCTCGTGAAGAGTCCGGCACTGAACACCCCGGGCCCGTCCACCTCC 5835  
Qy 5698 CGCAGCGCCCAACAGCGAGCGCCACCCAGTACACGAGCATCATCAGGCGGTG 5757  
Db 5836 CGCAGCGCCCAACAGCGAGCGCCACCCAGTACACGAGCATCATCAGGCGGTG 5895  
Qy 5758 GCTCTCAGCGCCAGCGCGCCGAGCGCCAGCCACCCCGGAGCGGACACACCCAC 5817  
Db 5896 GCTCTCAGCGCCAGCGCGCCGAGCGCCAGCCACCCCGGAGCGGACACACCCAC 5955  
Qy 5818 CGCTACCGCGAGGCGGACCGAGCTGCGCAGGACAGTCTCTGGCGCCCTCGAG 5877  
Db 5956 CGCTACCGCGAGGCGGACCGAGCTGCGCAGGACAGTCTCTGGCGCCCTCGAG 6015  
Qy 5878 CGAGAGAGTCTCCCGCGCGGTGTCTCAGCAGCGGAGAGCGGTCTCCCGGAGGCTG 5937  
Db 6016 CGAGAGAGTCTCCCGCGCGGTGTCTCAGCAGCGGAGAGCGGTCTCCCGGAGGCTG 6075  
Qy 5938 TTTGAAGACAGCAGCGGCGCGGTGTCTCGGAGCGGTGAGACCCCGGTGTCACG 5997  
Db 6076 TTTGAAGACAGCAGCGGCGCGGTGTCTCGGAGCGGTGAGACCCCGGTGTCACG 6135  
Qy 5998 GTCAACAGGAGAGGGCA 6017  
Db 6136 GTCAACAGGAGAGGGCA 6155

Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease; systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy; stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia; pain; alcoholism; transgenic.

Homo sapiens.

WO200226826-A2.

04-APR-2002.

27-SEP-2001; 2001WO-US042336.

27-SEP-2000; 2000US-0235631P.

27-SEP-2000; 2000US-0235633P.

27-SEP-2000; 2000US-0235808P.

27-SEP-2000; 2000US-0236064P.

27-SEP-2000; 2000US-0236065P.

27-SEP-2000; 2000US-0236066P.

28-SEP-2000; 2000US-0236135P.

05-OCT-2000; 2000US-0238321P.

06-OCT-2000; 2000US-0238396P.

06-OCT-2000; 2000US-0238399P.

16-MAR-2001; 2001US-0276667P.

31-MAY-2001; 2001US-0294823P.

12-JUL-2001; 2001US-0304868P.

26-SEP-2001; 2001US-00235631.

(CURA-) CURAGEN CORP.

Gerlach VL, Macdougall JR, Smithson G, Millet I, Stone D;

Gunther E, Ellerman K, Grosse WM, Alsobrook JP, Lepley DM;

Burgess CB, Padigar M, Kekuda R, Szytek KA, Leach MD, Shimkets RA;

WPI; 2002-499860/53.

P-PSDB; ABG78363.

Novel isolated NOVX polypeptides and polynucleotides homologous to

atractin, plexin, papin-like family of proteins, useful for treating

atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and

stroke.

Claim 8; Page 43-44; 308pp; English.

The invention discloses the isolated human polypeptides, and

polynucleotides encoding them, that have been designated NOVX. The

polypeptides, polynucleotides and antibodies are useful in treating or

preventing a NOVX-associated disorder which is cardiomyopathy,

atherosclerosis and diabetes in a human, where the disorder is related to

cell signal processing and metabolic pathway modulation. They can also be

used in determining the presence of, or predisposition to, a disease

associated with altered levels of the polypeptides and polynucleotides of

any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for

identifying an agent that binds to, or that modulates the expression or

activity of the polypeptide, for identifying an agent which is cellular

receptor or downstream effector, for treating or preventing a NOVX-

associated disorder and as a pharmaceutical composition comprising the

polypeptide, polynucleotide or the antibody. The polypeptides and

polynucleotides are useful in diagnostic applications (e.g. as a marker

for cancerous cells or tissue types) where their amounts are assessed, or

for the manufacture of a medicament (e.g. gene therapy) for treating or

preventing disorders or syndromes such as hypertension, congenital heart

defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,

Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,

haemophilia, dyslipidemia, haematopoietic diseases, scleroderma,

infertility, idiopathic thrombocytopenic purpura, graft versus host

diseases, Crohn's disease, multiple sclerosis, arthritis, autoimmune

diseases, systemic lupus erythematosus, asthma, psoriasis, ataxia,

allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar

ataxia, pain and alcoholism. They may also be used as immunogens to

produce antibodies specific for the invention, and as vaccines.

Transgenic cells containing a NOVX expressing construct are useful to

CC produce non-human transgenic animals for studying the function and/or  
CC activity of the NOVX proteins and for identifying and/or evaluating  
CC modulators of NOVX protein activity. Transgenic cells containing a NOVX  
CC expressing construct are useful to produce non-human transgenic animals  
CC for studying the function and/or activity of the NOVX proteins and for  
CC identifying and/or evaluating modulators of NOVX protein activity. The  
CC sequences presented in AB563431-AB563444 are the human NOV1-NOV8 cDNAs  
XX  
SQ Sequence 6139 BP; 1742 A; 1554 C; 1690 G; 1203 T; 0 U; 0 Other;  
Query Match 85.7%; Score 5631; DB 6; Length 6189;  
Best Local Similarity 95.3%; Pred. No. 0;  
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;  
QY 19 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 78  
DB 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 60  
QY 79 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTTATGACTCAA 138  
DB 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTTATGACTCAA 120  
QY 139 CAGCAGATGTCCTCTTTCCGAGAGGATATTAGATGCCCTCTTTGTTCTTTTGAA 198  
DB 121 CAGCAGATGTCCTCTTTCCGAGAGGATATTAGATGCCCTCTTTGTTCTTTTGAA 180  
QY 199 GAATGAGTCAGCTGCTCTGATGAGATTAAAGCAGCTGAGCACTTCTCCGGAGTAT 258  
DB 181 GAATGAGTCAGCTGCTCTGATGAGATTAAAGCAGCTGAGCACTTCTCCGGAGTAT 240  
QY 259 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGAGCTTCGAAGTCAGA 318  
DB 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGAGCTTCGAAGTCAGA 300  
QY 319 AGTCTTGATGTTGGTCACTTTGCTGAAGTCAGTGGTGAAGAGAAAGCCGG 378  
DB 301 AGTCTTGATGTTGGTCACTTTGCTGAAGTCAGTGGTGAAGAGAAAGCCGG 360  
QY 379 GACATCTATCTATGAAGTGAAGAGAGAGCTTTATTTGGCCAGGAGAGGTTTCA 438  
DB 361 GACATCTATCTATGAAGTGAAGAGAGAGCTTTATTTGGCCAGGAGAGGTTTCA 420  
QY 439 TTTTGGAGAGAGCGGAAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 498  
DB 421 TTTTGGAGAGAGCGGAAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 480  
QY 499 CAGTATGCCCTTTCAGGACAAATATCACTTTATCTGATGGAGGAATATCAGCTGGAGG 558  
DB 481 CAGTATGCCCTTTCAGGACAAATATCACTTTATCTGATGGAGGAATATCAGCTGGAGG 540  
QY 559 GACTTGTCTGCTACTTTTGAATAGATAGGACCACTTAGATGAAACCTGTATACAGTTT 618  
DB 541 GACTTGTCTGCTACTTTTGAATAGATAGGACCACTTAGATGAAACCTGTATACAGTTT 600  
QY 619 TACCTAGCTGAGCTGATTTGGCTGTTACAGGTTTCATCTGATGGATACCTGTATCGA 678  
DB 601 TACCTAGCTGAGCTGATTTGGCTGTTACAGGTTTCATCTGATGGATACCTGTATCGG 660  
QY 679 GACATCAAGCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTATTTT 738  
DB 661 GACATCAAGCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTATTTT 720  
QY 739 GGATCTGCCGGAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTTGGGACC 798  
DB 721 GGATCTGCCGGAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTTGGGACC 777  
QY 799 CCAGATTACATGCTCTCTGAAGTGTGACTGTGATGAACGGGGATGGAAGGCACTTAC 858  
DB 778 CCAGATTACATGCTCTCTGAAGTGTGACTGTGATGAACGGGGATGGAAGGCACTTAC 837  
QY 859 GGCCTGGATGTGACTGTTGGTGTGAGTGGCGGTGATTTGCCCTATGAGATGATTTATGGGAGA 918  
DB 838 GGCCTGGATGTGACTGTTGGTGTGAGTGGCGGTGATTTGCCCTATGAGATGATTTATGGGAGA 897

QY 919 TCCCTCTTCGAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTAATTAATTTCCAGGG 978  
DB 898 TCCCTCTTCGAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTAATTAATTTCCAGGG 957  
QY 979 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1038  
DB 958 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1017  
QY 1039 TTTTGTTCGCGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTTGTCTCCCATCTCTTC 1098  
DB 1018 TTTTGTTCGCGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTTGTCTCCCATCTCTTC 1077  
QY 1099 TCTAAATTTGACTCGAACAACATTCGTAACCTCTCTCCCTCCCTCTCGTTCACCCCTCAAG 1158  
DB 1078 TCTAAATTTGACTCGAACAACATTCGTAACCTCTCTCCCTCCCTCTCGTTCACCCCTCAAG 1137  
QY 1159 TCTGACGATGACACTCCAAATTTTGTGAACACAGAGAGAAATTCGTTGGTTCATCCTCT 1218  
DB 1138 TCTGACGATGACACTCCAAATTTTGTGAACACAGAGAGAAATTCGTTGGTTCATCCTCT 1197  
QY 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGCTTTGTGGGTTTCG 1278  
DB 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGCTTTGTGGGTTTCG 1257  
QY 1279 TACAGAGGCACTGGGATTTTGTGTAGATCTGAGTCTGTTGTGCGGCTCTGGACTCC 1338  
DB 1258 TACAGAGGCACTGGGATTTTGTGTAGATCTGAGTCTGTTGTGCGGCTCTGGACTCC 1317  
QY 1339 CCTCCAGACTAGCTCCATCGAAAGAAATCTTCTCATAAAAGCAAGAGCTACAAGAC 1398  
DB 1318 CCTCCAGACTAGCTCCATCGAAAGAAATCTTCTCATAAAAGCAAGAGCTACAAGAC 1377  
QY 1399 TCTCAGAGCAAGTGTCAAGATGGAGCAGAAATGACCGGTTTACATCGAGAGTGTCA 1458  
DB 1378 TCTCAGAGCAAGTGTCAAGATGGAGCAGAAATGACCGGTTTACATCGAGAGTGTCA 1437  
QY 1459 GAGTGGAGGCTGTCTTAGTTCAGAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518  
DB 1438 GAGTGGAGGCTGTCTTAGTTCAGAGAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1497  
QY 1519 TCCCTCTTCGAGCAGGACCTTGTACCTACATCACAGAAATGCAAGTGAAGTGAAGTGAAGT 1578  
DB 1498 TCCCTCTTCGAGCAGGACCTTGTACCTACATCACAGAAATGCAAGTGAAGTGAAGTGAAGT 1557  
QY 1579 TTGAGCAAGCACGGATGGAGTGTCTCCAGAGGAGTGAAGGCTGCAAAAGCACTGCAAGTTCCTCAT 1638  
DB 1558 TTGAGCAAGCACGGATGGAGTGTCTCCAGAGGAGTGAAGGCTGCAAAAGCACTGCAAGTTCCTCAT 1617  
QY 1639 GATATCAGAGCAGAGCCCGAAGCTCCAAAGAAATCAAAGCAGAGGATACAGGCTCAA 1698  
DB 1618 GATATCAGAGCAGAGCCCGAAGCTCCAAAGAAATCAAAGCAGAGGATACAGGCTCAA 1677  
QY 1699 GTGGAAGAAATGAGTTGATGATGAATCAGTTCGAGAGGATCTTGTCTCAGCAAGAGA 1758  
DB 1678 GTGGAAGAAATGAGTTGATGATGAATCAGTTCGAGAGGAGATCTTGTCTCAGCAAGAGA 1737  
QY 1759 CGGAGTGAATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1818  
DB 1738 CGGAGTGAATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG 1797  
QY 1819 CGGAAGCGCAGAAATGTCAGCTTAACCTTGAAGGCTTAAGGATCAAGGGAAGCTTGA 1878  
DB 1798 CGGAAGCGCAGAAATGTCAGCTTAACCTTGAAGGCTTAAGGATCAAGGGAAGCTTGA 1857  
QY 1879 GTGGGAGAAATGCGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938  
DB 1858 GTGGGAGAAATGCGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1917  
QY 1939 CTCCAGAGAACTGGAGAGGCTG----- 1963  
DB 1918 CTCCAGAGAACTGGAGAGGCTGTAAAGCCAGACCGAGGCCACCGAGCTGCTGCAG 1977

QY 1964 -----CAGAAGAGCGAGCCGAGAGGGAGCTGGAGAAAGCTGCAGAACCGAGAG 2010  
Db 1978 AATATCCGCGACGCAAAAGGAGCGAGCCGAGAGGGAGCTGGAGAAAGCTGCAGAACCGAGAG 2037  
QY 2011 GATTCTTCTGAAGGCATCAAAAGAAAGCTGGTGGAGCTGAGAAAGCGCGCATTTCTTG 2070  
Db 2038 GATTCTTCTGAAGGCATCAAAAGAAAGCTGGTGGAGCTGAGAAAGCGCGCATTTCTTG 2097  
QY 2071 GAGAACAAAGGTAAAGAGATTAGAGACCATGAGAGCCTGAGAGCCTAGAGAAACAGACTGAAGATGAC 2130  
Db 2098 GAGAACAAAGGTAAAGAGATTAGAGACCATGAGAGCCTGAGAGCCTAGAGAAACAGACTGAAGATGAC 2157  
QY 2131 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCCAGAG 2190  
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCCAGAG 2217  
QY 2191 AAACATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAAGTCCACCTGAAACAGAAAGAG 2250  
Db 2218 AAACATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAAAGTCCACCTGAAACAGAAAGAG 2277  
QY 2251 CAGCACTATGAGGAAGAAGATTAAAGTTTGGACAATCAGATAAAGAAAGACCTGCTGAC 2310  
Db 2278 CAGCACTATGAGGAAGAAGATTAAAGTTTGGACAATCAGATAAAGAAAGACCTGCTGAC 2337  
QY 2311 AAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGCGCCCATGAGAAAGGCCAAA 2370  
Db 2338 AAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGGCGCCCATGAGAAAGGCCAAA 2397  
QY 2371 ATTCTCAGCGAACAGAAAGCGATGATCAATGCTATGGATTCAGATCCAGATCAGATCCCTGGAA 2430  
Db 2398 ATTCTCAGCGAACAGAAAGCGATGATCAATGCTATGGATTCAGATCCAGATCAGATCCCTGGAA 2457  
QY 2431 CAGAGATTGTGGAGTGTGAGCGCAATAAATTCGACGAAATAGCAGTCTTTTACC 2490  
Db 2458 CAGAGATTGTGGAGTGTGAGCGCAATAAATTCGACGAAATAGCAGTCTTTTACC 2517  
QY 2491 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2550  
Db 2518 CAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577  
QY 2551 CTGGAGACACAGCTGGGAAGTTGGAGCCGAGAACCGAATCTGGAGGAGCAGCTGGAG 2610  
Db 2578 CTGGAGACACAGCTGGGAAGTTGGAGCCGAGAACCGAATCTGGAGGAGCAGCTGGAG 2637  
QY 2611 AAGATCAGCCACCAAGACACACAGTGCAGAGAAATCGGCTGCTGGAATCTGGAGCAAGATTG 2670  
Db 2638 AAGATCAGCCACCAAGACACACAGTGCAGAGAAATCGGCTGCTGGAATCTGGAGCAAGATTG 2697  
QY 2671 CGGAGGTCAAGTCTAGAGACAGAGGAGAGAAATCTGGAGCTCAAGGCGCAGCTCACAGAG 2730  
Db 2698 CGGAGGTCAAGTCTAGAGACAGAGGAGAGAAATCTGGAGCTCAAGGCGCAGCTCACAGAG 2757  
QY 2731 CTACAGCTCTCCTGCGAGGCGCGAGTCCACAGTTGACAGCCCTGCAGGCTGCACGGCGG 2790  
Db 2758 CTACAGCTCTCCTGCGAGGCGCGAGTCCACAGTTGACAGCCCTGCAGGCTGCACGGCGG 2817  
QY 2791 GGCCTGGAGAGCGAGTTTCCGAGGCGAAGACAGAGCTGGAAGAGACACACAGCAAGCT 2850  
Db 2818 GGCCTGGAGAGCGAGTTTCCGAGGCGAAGACAGAGCTGGAAGAGACACACAGCAAGCT 2877  
QY 2851 GAAGAGGATCCAGGCACTCAGGCGACATAGATGAAATCCAGCGCAAAATTTGATGCT 2910  
Db 2878 GAAGAGGATCCAGGCACTCAGGCGACATAGATGAAATCCAGCGCAAAATTTGATGCT 2937  
QY 2911 CTTTCGTAAACAGTGTACTGTAATCAAGACCTGGAGGAGCAGCTAAACAGCTGACCGGAG 2970  
Db 2938 CTTTCGTAAACAGTGTACTGTAATCAAGACCTGGAGGAGCAGCTAAACAGCTGACCGGAG 2997  
QY 2971 GACACGCTGAACCTAAACACCAAACTTCTACTTGTTCAAAACAACTGATGAGGCTTCT 3030  
Db 2998 GACACGCTGAACCTAAACACCAAACTTCTACTTGTTCAAAACAACTGATGAGGCTTCT 3057  
QY 3031 GGCGCCAAACGACGAGATTGTACAACTGCGAAGTGAAGTGGACCATCTCCGCGCGGAGATC 3090

Db 3058 GGCCCCAAACGACGAGATTGTACAACTCGAAGTGAAGTGAACCACTCTCCGCGGAGATC 3117  
QY 3091 ACGGAAACGAGAGATGAGCTTACCAGCCAGAACCAACGATGGAGGCTCTGAAACACCAAG 3150  
Db 3118 ACGGAAACGAGAGATGAGCTTACCAGCCAGAACCAACGATGGAGGCTCTGAAACACCAAG 3177  
QY 3151 TGCACCAATGCTGGAGGAACAGAGTATGATTTGGAGGCCCTTAAACGATGAGCTGCTAGAA 3210  
Db 3178 TGCACCAATGCTGGAGGAACAGAGTATGATTTGGAGGCCCTTAAACGATGAGCTGCTAGAA 3237  
QY 3211 AAAGAGCGGAGTGGAGGCGCTGGAGAGCGCTCCTGGGTGATGAGAAATCCCAGTTTGGAG 3270  
Db 3238 AAAGAGCGGAGTGGAGGCGCTGGAGAGCGCTCCTGGGTGATGAGAAATCCCAGTTTGGAG 3297  
QY 3271 TGTCCGTTTCGAGAGCTGCAGAGAAATCTGGACACCCAGAAAAACAGACAGGGCGAGAGCC 3330  
Db 3298 TGTCCGTTTCGAGAGCTGCAGAGAAATCTGGACACCCAGAAAAACAGACAGGGCGAGAGCC 3357  
QY 3331 GATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGAGTGAAGGAGCACAAGGCT 3390  
Db 3358 GATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGAGTGAAGGAGCACAAGGCT 3417  
QY 3391 GAGATTCTCGTCTGCAGCAGGCTCTCAAAGAGCAGAAAGCTGGAAGGCCGAGAGCTCTCT 3450  
Db 3418 GAGATTCTCGTCTGCAGCAGGCTCTCAAAGAGCAGAAAGCTGGAAGGCCGAGAGCTCTCT 3477  
QY 3451 GACAACTCAATGACCTGGAGAGAACATGCTATGCTTGAATGAATGCCCGAGGCTTA 3510  
Db 3478 GACAACTCAATGACCTGGAGAGAACATGCTATGCTTGAATGAATGCCCGAGGCTTA 3537  
QY 3511 CAGCAGAAAGCTGGAGATGAAACGAGAGCTCAAACAGAGGCTTCTGGAAGAGCAGAACCAAA 3570  
Db 3538 CAGCAGAAAGCTGGAGATGAAACGAGAGCTCAAACAGAGGCTTCTGGAAGAGCAGAACCAAA 3597  
QY 3571 TTAAGCAGCAGATGGAACCTGCAGAAAAATCAATTTCCGTCTGACTCAAGAGCTGCAA 3630  
Db 3598 TTAAGCAGCAGATGGAACCTGCAGAAAAATCAATTTCCGTCTGACTCAAGAGCTGCAA 3657  
QY 3631 GAAGCTTAGATCGGCTGATCTACTCAAGACAGAAAGAGTGAATGGAGTATCAGCTG 3690  
Db 3658 GAAGCTTAGATCGGCTGATCTACTCAAGACAGAAAGAGTGAATGGAGTATCAGCTG 3717  
QY 3691 GAAACCAATTCAGGTTCTTATTTCTATGAAAGGTGAAATGGAAGGCATTTTCTCAA 3750  
Db 3718 GAAACCAATTCAGGTTCTTATTTCTATGAAAGGTGAAATGGAAGGCATTTTCTCAA 3777  
QY 3751 CAAACCAACTCATTTGATTTCTGCAAGCCAAATGCAACCACTGCTTAAAGAAAG 3810  
Db 3778 CAAACCAACTCATTTGATTTCTGCAAGCCAAATGCAACCACTGCTTAAAGAAAG - - 3835  
QY 3811 GGTTTATTTAGTCGACGGAAGAGGACCTGCTTTACCCACACAGGTTCTCTGCAATAC 3870  
Db 3836 -----AGGTGCTCTGCAATAC 3852  
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGCTGTGCAAGAGCTAGAGGAGCC 3930  
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGCTGTGCAAGAGCTAGAGGAGCC 3912  
QY 3931 CTTTCAGAAAGACCCGATCTGAGCTCCGCTCCGCGGAGGAGCTGCCACCGCAAGCA 3990  
Db 3913 CTTTCAGAAAGACCCGATCTGAGCTCCGCTCCGCGGAGGAGCTGCCACCGCAAGCA 3972  
QY 3991 ACGGACCAACCAACCCATCCAGCCAGCCACCGCGAGGAGCAGATGCGCATGTCGCC 4050  
Db 3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGCGAGGAGCAGATGCGCATGTCGCC 4032  
QY 4051 ATCGTGGGTGCGCAGAGACCGCCAGTGCATGAGCTGCTGCGCCCGCCATCCAGC 4110  
Db 4033 ATCGTGGGTGCGCAGAGACCGCCAGTGCATGAGCTGCTGCGCCCGCCATCCAGC 4092  
QY 4111 CGCAGAAAGAGGCTTCAACTCCAGAGGAATTTAGTCGGGCTCTTAAAGAACGATGCAC 4170

4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAAACGATGCAC 4152  
4171 CACAATATTTCTTCAACGATTCAACGTAGGACTGAAATCGGAGCCCAAAAGTGTGCTGTG 4230  
4153 CACAATATTTCTTCAACGATTCAACGTAGGACTGAAATCGGAGCCCAAAAGTGTGCTGTG 4212  
4231 TGTCTGATACCGTGCACATTTGGACGCCAGCCATCCAAATGTCTCGAATGTCTCAGGTGATG 4290  
4213 TGTCTGATACCGTGCACATTTGGACGCCAGCCATCCAAATGTCTAGAATGTCTCAGGTGATG 4272  
4291 TGTACCCCCAAGTGTCTCCAGTGTGTCGACGACACCTGCGGCTTGGCTGCTGATATGCC 4350  
4273 TGTACCCCCAAGTGTCTCCAGTGTGTCGACGACACCTGCGGCTTGGCTGCTGATATGCC 4332  
4351 ACACACTTACCCGAGGCTTCTCCGCTGACAAATGAATCTCCCGAGGTCTCCAGACCAAG 4410  
4333 ACACACTTACCCGAGGCTTCTCCGCTGACAAATGAATCTCCCGAGGTCTCCAGACCAAG 4392  
4411 GAGCCACGACGCTTGCACCTGGAAGGTGATGAAGTGCACGAGTCCAGGAATACAAACGA 4470  
4393 GAGCCACGACGCTTGCACCTGGAAGGTGATGAAGTGCACGAGTCCAGGAATACAAACGA 4452  
4471 GGACAGCAGGCTGGACAGGAGTACATTTCTCGAGGAGTCAAAAGTCTCAATTTAT 4530  
4453 GGACAGCAGGCTGGACAGGAGTACATTTCTCGAGGAGTCAAAAGTCTCAATTTAT 4512  
4531 GACAAATGAAGCCAGAGAGCTGGACAGGCGCGTGGAAAGATTTGAGTGTGCCCTTCCC 4590  
4513 GACAAATGAAGCCAGAGAGCTGGACAGGCGCGTGGAAAGATTTGAGTGTGCCCTTCCC 4572  
4591 GACGGGATGTATCTATTCATGTGCGGTGTGCTTCGAACTCCAAATACAGCCAAA 4650  
4573 GACGGGATGTATCTATTCATGTGCGGTGTGCTTCGAACTCCAAATACAGCCAAA 4632  
4651 GCA----- 4653  
4633 GCAGATGTCCTACATACATCTGAAGATGGAATCTCACCCGACACACCTGCTGGCCCGG 4692  
4654 ----- 4653  
4693 AGAACCCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4752  
4654 -----GAAAAAGCAGAGCTGATGCTAAACTG 4680  
4753 GAATCAGTTCTCCAGGTGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATGCTAAACTG 4812  
4681 CTTGGAAACTCTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGTG 4740  
4813 CTTGGAAACTCTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAACCTGCACGTG 4872  
4741 CCTTCAGTGACAGGTGTGTGGTGGCCACCGAGGAGGCTCTACGCCCTGATGTC 4800  
4873 CCTTCAGTGACAGGTGTGTGGTGGCCACCGAGGAGGCTCTACGCCCTGATGTC 4932  
4801 TTGAAAACTCTCCCTAACCCATGTCCAGGAAATGGAGAGTCTTCCAAATTTATATTATC 4860  
4933 TTGAAAACTCTCCCTAACCCATGTCCAGGAAATGGAGAGTCTTCCAAATTTATATTATC 4992  
4861 AAGGACTGGAGAGTCTCATGTATAGCAGGAGAGCGGGCACTGTCTTGTGGAC 4920  
4993 AAGGACTGGAGAGTCTCATGTATAGCAGGAGAGCGGGCACTGTCTTGTGGAC 5052  
4921 GTGAGAAAGTGAACAGTCTCCCTGGCCAGTCCCACTGCCCTGCCCGCAGCTCTCA 4980  
5053 GTGAGAAAGTGAACAGTCTCCCTGGCCAGTCCCACTGCCCTGCCCGCAGCTCTCA 5112  
4981 CCAGCATTTTGAAGTGTCAAGGCTGCGCACTTTTGGGGCAGGCAAGATTGAGAAC 5040  
5113 CCAGCATTTTGAAGTGTCAAGGCTGCGCACTTTTGGGGCAGGCAAGATTGAGAAC 5172  
5041 GGGCTCTGATCTGTGACGACATGCCAGCAAGTCTCATTTCTCGCTTACAAACCAAAAC 5100  
5173 GGGCTCTGATCTGTGACGACATGCCAGCAAGTCTCATTTCTCGCTTACAAACCAAAAC 5232

5101 CTCAGCAATATCTGCATCCCGGAAAGAGATAGAGACCTCAGAGCCCTCGAGCTGTATCCAC 5160  
5233 CTCAGCAATATCTGCATCCCGGAAAGAGATAGAGACCTCAGAGCCCTCGAGCTGTATCCAC 5292  
5161 TTCACCAATACAGTATCTCTCAATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5220  
5293 TTCACCAATACAGTATCTCTCAATTTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5352  
5221 TACAGCTCGAGGAATTCCTGGATAAAGATGACATTTCTTGGCACCTGTGTGTTTGGC 5280  
5353 TACAGCTCGAGGAATTCCTGGATAAAGATGACATTTCTTGGCACCTGTGTGTTTGGC 5412  
5281 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGCGCAGGAG 5340  
5413 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGCGCAGGAG 5472  
5341 GAGTACTTGTGTGTTTCCACGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5400  
5473 GAGTACTTGTGTGTTTCCACGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5532  
5401 CGCACAGACGATCTCAAGTGGAGTGGCTTACCTTTGGCCTTTTGCTTACAGAGAACCTAT 5460  
5533 CGCACAGACGATCTCAAGTGGAGTGGCTTACCTTTGGCCTTTTGCTTACAGAGAACCTAT 5592  
5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTTGAGATCCAGGACCGCTCTCAGCA 5520  
5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTTGAGATCCAGGACCGCTCTCAGCA 5652  
5521 GGGACCCCTCGCCAGAGTACTTGGACATCCGAAACCGCGCTACCTGGGCCCTTGCCTAT 5580  
5553 GGGACCCCTCGCCAGAGTACTTGGACATCCGAAACCGCGCTACCTGGGCCCTTGCCTAT 5712  
5581 TCCTCAGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGTTCATTGTCTGC 5640  
5713 TCCTCAGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGTTCATTGTCTGC 5772  
5641 AAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACACCGGGCCCGTCCACCTCCCG 5700  
5773 AAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACACCGGGCCCGTCCACCTCCCG 5832  
5701 AGCAGCCCCAACAGCAGGCGCCACCCACGTAACAAGAGCACATCAACAAGCGCTGGCC 5760  
5833 AGCAGCCCCAACAGCAGGCGCCACCCACGTAACAAGAGCACATCAACAAGCGCTGGCC 5892  
5761 TCCAGCCAGGCGCGCCGAGAGGCGCCAGCCACCGGAGAGCCAGCACACCCCGCC 5820  
5893 TCCAGCCAGGCGCGCCGAGAGGCGCCAGCCACCGGAGAGCCAGCACACCCCGCC 5952  
5821 TACCGCAGGCGCGGAGCCGAGTGCAGGAGCAAGTCTCTTGGCGCCCGCTGGAGCGA 5880  
5953 TACCGCAGGCGCGGAGCCGAGTGCAGGAGCAAGTCTCTTGGCGCCCGCTGGAGCGA 6012  
5881 GAGAAGTCCCGCGCCGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGCGTGT 5940  
6013 GAGAAGTCCCGCGCCGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGCGTGT 6072  
5941 GAAGACAGCAGGCGCGGCTGCTCGGAGAGCGGTGAGGACCCCGCTGTCCAGGTG 6000  
6073 GAAGACAGCAGGCGCGGCTGCTCGGAGAGCGGTGAGGACCCCGCTGTCCAGGTG 6132  
6001 AACAGGAAAGAGGCGAGAGTGTCTCAAGTTTTCACGGTTTAACTGTACCT 6055  
6133 AACAGGTTGGGACCACTTTCAGTATAAATCTCAGCCAGAAAAAACCACTCCT 6187

RESULT 9  
ADA05641  
ID ADA05641 standard; cDNA; 6189 BP.  
XX  
AC ADA05641;  
XX  
DT 06-NOV-2003 (first entry)

XX DE Human NOV1a encoding cDNA SEQ ID NO:1.

XX KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

XX KW immunomodulator; cytostatic; nootropic; neuroprotective;

XX KW antiparkinsonian; antilipemic; gene therapy; human disease;

XX KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

XX KW immune disorder; haematopoietic disorder; dyslipidemia; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..6162

XX FT /\*tag= "NOV1a"

XX FT /product= "NOV1a"

XX PN WO2003029424-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031373.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327433P.

XX PR 05-OCT-2001; 2001US-0327449P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 17-APR-2002; 2002US-0373260P.

XX PR 19-APR-2002; 2002US-0373815P.

XX PR 19-APR-2002; 2002US-0373817P.

XX PR 19-APR-2002; 2002US-0373826P.

XX PR 19-APR-2002; 2002US-0373884P.

XX PR 22-APR-2002; 2002US-0374977P.

XX PR 16-MAY-2002; 2002US-0381037P.

XX PR 16-MAY-2002; 2002US-0381038P.

XX PR 17-MAY-2002; 2002US-0381042P.

XX PR 17-MAY-2002; 2002US-0381642P.

XX PR 28-MAY-2002; 2002US-0383656P.

XX PR 29-MAY-2002; 2002US-0383831P.

XX PR 25-JUN-2002; 2002US-0391335P.

XX PR 01-OCT-2002; 2002US-00262511.

XX PA (CURA-) CURAGEN CORP.

XX PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

XX PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

XX PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

XX PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;

XX PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

XX PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX DR WPI; 2003-381626/36.

XX DR P-PSDB; ADA05642.

XX XX New NOVX polypeptides and nucleic acids, useful for diagnosing,

XX PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,

XX PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

XX PT pharmacogenomics.

XX XX Claim 20; Page 98-99; 586pp; English.

XX PS

XX XX

CC The present invention describes NOVX proteins, where X can be 1 to 55

CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide

CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector

CC comprising the nucleic acid molecule described above; (5) a cell

CC comprising the above vector; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to

CC a disease associated with altered levels of expression of the above

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a

CC method of identifying an agent that binds to the polypeptide described

CC above; (10) a method for identifying a potential therapeutic agent for

CC use in treating a pathology that is related to an aberrant expression or

CC aberrant physiological interactions of the polypeptide; (11) a method of

CC screening for a modulator of activity or of latency or predisposition to

CC a pathology associated with the polypeptide; (12) a method for modulating

CC the activity of the polypeptide described above; (13) methods of treating

CC or preventing a pathology associated with the above polypeptide in a

CC mammal; and (14) a method for producing the above polypeptide. NOVX

CC sequences have antidiabetic, anorectic, antibacterial, virucide,

CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian

CC and antilipemic activities, and can be used in gene therapy. The

CC polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease. The polypeptide or the nucleic

CC acid molecule may be used to diagnose, treat or prevent metabolic

CC disorders such as diabetes or obesity, infections, cachexia, cancer,

CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's

CC disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation

CC probes, in chromosome mapping, tissue typing, preventive medicine and

CC pharmacogenomics. The present sequence encodes a human NOVX protein from

XX the present invention.

SQ Sequence 6189 BP; 1742 A; 1554 C; 1690 G; 1203 T; 0 U; 0 Other;

Query Match 85.7%; Score 5631; DB 7; Length 6189;

Best Local Similarity 95.3%; Pred. No. 0;

Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;

QY 19 ATGTTGAAGTTCAATATGGAGCGCGGAATCCTTTGGATGCTGGTGTGCTGAACCCATT 78

DB 1 ATGTTGAAGTTCAATATGGAGCGCGGAATCCTTTGGATGCTGGTGTGCTGAACCCATT 60

QY 79 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCTTTATGACTCAA 138

DB 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCTTTATGACTCAA 120

QY 139 CAGCAGATGTCCTCTTCCCGAGAGGGATATTAGATGCCCTCTTGTCTCTTTTGA 198

DB 121 CAGCAGATGTCCTCTTCCCGAGAGGGATATTAGATGCCCTCTTGTCTCTTTTGA 180

QY 199 GAATGCAGTCAGCCTGCTCTGATGAAGATTAAAGCACGTGAGCAACTTTGTCGGAAGTAT 258

DB 181 GAATGCAGTCAGCCTGCTCTGATGAAGATTAAAGCACGTGAGCAACTTTGTCGGAAGTAT 240

QY 259 TCCGACACCATAGCTAGTGTACAGGACTCCAGCTTCGCGCAAGGACTTCGAAGTCAGA 318

DB 241 TCCGACACCATAGCTAGTGTACAGGACTCCAGCTTCGCGCAAGGACTTCGAAGTCAGA 300

QY 319 AGTCTTGTAGGTTCTGTTGCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAGCAACCGGG 378

DB 301 AGTCTTGTAGGTTCTGTTGCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAGCAACCGGG 360

QY 379 GACATCTATGCTATGAAGTATGAAGAAGGCTTTATGGCCCGAGGACAGGTTTCA 438

DB 361 GACATCTATGCTATGAAGTATGAAGAAGGCTTTATGGCCCGAGGACAGGTTTCA 420

QY 439 TTTTGTGAGGAGCGGACATATTATTCGAGACCAAGCCCGTGTGATCCCCCAATTA 498

DB 421 TTTTGTGAGGAGCGGACATATTATTCGAGACCAAGCCCGTGTGATCCCCCAATTA 480



Db 2638 AAGATCAGCCACCAAGACACACAGTGCACAAAGAAATCGGCTGCTGGAATCGGAGTGAACAAGATTG 2697  
Qy 2671 CGGGAGTCAGTCTAGACGACGAGGAGCAGAACTGGGCTCAAGCGCCAGCTCCACAG 2730  
Db 2698 CGGGAGTCAGTCTAGACGACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCCACAG 2757  
Qy 2731 CTACAGCTCTCCCTGACGAGCGCGAGTCACAGTTGACAGCCCTGACGAGCTGCACGGCG 2790  
Db 2758 CTACAGCTCTCCCTGACGAGCGCGAGTCACAGTTGACAGCCCTGACGAGCTGCACGGCG 2817  
Qy 2791 GCCCTGGAGCGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGAGCT 2850  
Db 2818 GCCCTGGAGCGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGAGCT 2877  
Qy 2851 GAAGAGAGATCCAGGCACTCCAGGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2910  
Db 2878 GAAGAGAGATCCAGGCACTCCAGGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2937  
Qy 2911 CTTCTGTAACAGCTGACTGTAATCAAGACCTGGAGAGCAGCTAAACAGCTGACCGAG 2970  
Db 2938 CTTCTGTAACAGCTGTAATCAAGACCTGGAGAGCAGCTAAACAGCTGACCGAG 2997  
Qy 2971 GACACGCTGAACTCAACCAACAAACTTCTACTTCTCAAAACAATCCGATGAGGCTTCT 3030  
Db 2998 GACACGCTGAACTCAACCAACAAACTTCTACTTCTCAAAACAATCCGATGAGGCTTCT 3057  
Qy 3031 GCGCCAAACGACGAGATTGTAACTGCGAAGTGAAGTGCACATCTCCGCCGGGAGATC 3090  
Db 3058 GCGCCAAACGACGAGATTGTAACTGCGAAGTGAAGTGCACATCTCCGCCGGGAGATC 3117  
Qy 3091 ACGGACGAGATGAGCTTACCAGCCAGAGACCAACGATGGAGCTCTGAAGACCCAG 3150  
Db 3118 ACGGACGAGATGAGCTTACCAGCCAGAGACCAACGATGGAGCTCTGAAGACCCAG 3177  
Qy 3151 TGCACCATCTCGAGGAAACAGGTCATGATTTGGAGCCCTAAACGATGAGCTGTAGAA 3210  
Db 3178 TGCACCATCTCGAGGAAACAGGTCATGATTTGGAGCCCTAAACGATGAGCTGTAGAA 3237  
Qy 3211 AAAGACGAGTCGGAGGCTGGAGAGCGTCTGGGTCATGAGAAATCCAGTTGAG 3270  
Db 3238 AAAGACGAGTCGGAGGCTGGAGAGCGTCTGGGTCATGAGAAATCCAGTTGAG 3297  
Qy 3271 TGTCCGGTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGAGGCGAGAGCC 3330  
Db 3298 TGTCCGGTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGAGGCGAGAGCC 3357  
Qy 3331 GATCAGCGATCACGAGTCTGCCAGTGGTGGAGCTGGCAGTGGCAGTGAAGAGCACAGGCT 3390  
Db 3358 GATCAGCGATCACGAGTCTGCCAGTGGTGGAGCTGGCAGTGAAGAGCACAGGCT 3417  
Qy 3391 GAGATTCTCGCTCGACGAGGCTCTCAAGAGCAGAACTGAAGCCGAGAGCCTCTCT 3450  
Db 3418 GAGATTCTCGCTCGACGAGGCTCTCAAGAGCAGAACTGAAGCCGAGAGCCTCTCT 3477  
Qy 3451 GACAAGCTCAATGACCTGGAGAGAAAGCATGCTATCTTTGAAATGAATGCCGAAAGCTTTA 3510  
Db 3478 GACAAGCTCAATGACCTGGAGAGAAAGCATGCTATCTTTGAAATGAATGCCGAAAGCTTTA 3537  
Qy 3511 CAGCAGAAGCTGAGACTGAACAGAGCTCAACAGAGGCTTCTGAAAGAGCAAGCCAAA 3570  
Db 3538 CAGCAGAAGCTGAGACTGAACAGAGCTCAACAGAGGCTTCTGAAAGAGCAAGCCAAA 3597  
Qy 3571 TTACAGCAGCAGATGGACCTGACAGAAATACATTTCCGTCGACTCAAGGACTGCAA 3630  
Db 3598 TTACAGCAGCAGATGGACCTGACAGAAATACATTTCCGTCGACTCAAGGACTGCAA 3657  
Qy 3631 GAAGCTCTAGATCGGCTGATCTACTGAAACAGAAAGAGTGAATTTGAGTATCAGCTG 3690  
Db 3658 GAAGCTCTAGATCGGCTGATCTACTGAAACAGAAAGAGTGAATTTGAGTATCAGCTG 3717  
Qy 3691 GAAAAATTCAGCTTCTTATCTCATGAAAGGTGAATGAAGGCACTTATTTCTCAA 3750  
Db 3718 GAAAAATTCAGCTTCTTATCTCATGAAAGGTGAATGAAGGCACTTATTTCTCAA 3777

Qy 3751 CAAACCAAACTCATTGATTTTCTGAAAGCCAAAATGACCAACCTGCTAAAAAGAAAAAG 3810  
Db 3778 CAAACCAAACTCATTGATTTTCTGAAAGCCAAAATGACCAACCTGCTAAAAAGAAAA-- 3835  
Qy 3811 GGTTTATTTAGTCAGCGAAAGAGGACCTTGCTTTTACCACACAGGTTCTCTCAGTAC 3870  
Db 3836 -----AGGTGCTCTCAGTAC 3852  
Qy 3871 AATGAGCTGAAGCTGGCCCTGGAGAAAGGAAAGCTCGCTGTGCAGAGCTAGAGAAAGCC 3930  
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAAAGGAAAGCTCGCTGTGCAGAGCTAGAGAAAGCC 3912  
Qy 3931 CTTCAGAAAGCCCAATCGAGCTCCGTCGCCCGGAGGAAGCTGCCACCGCAAGCA 3990  
Db 3913 CTTCAGAAAGCCCAATCGAGCTCCGTCGCCCGGAGGAAGCTGCCACCGCAAGCA 3972  
Qy 3991 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGCAGATCGCATCTCGCC 4050  
Db 3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGCAGATCGCATCTCTGCC 4032  
Qy 4051 ATCGTGGGTCGCCAGAGCACCAGCCAGTGCCATGAGCTGTGGCCCGCCATCCAGC 4110  
Db 4033 ATCGTGGGTCGCCAGAGCACCAGCCAGTGCCATGAGCTGTGGCCCGCCATCCAGC 4092  
Qy 4111 CGGAAAGAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGTCTTAAGGAACGATGCAC 4170  
Db 4093 CGGAAAGAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGTCTTAAGGAACGATGCAC 4152  
Qy 4171 CACAATTTCTCTCACCAGCTCAAGTAGGACTGAACATGCGAGGACCAAAAGTGTGCTGTG 4230  
Db 4153 CACAATTTCTCTCACCAGCTCAAGTAGGACTGAACATGCGAGGACCAAAAGTGTGCTGTG 4212  
Qy 4231 TGTCTGATACCGTGCACTTTTGGACGCGAGGATCCAAAATGCTCAGAAATGTCAGGTGATG 4290  
Db 4213 TGTCTGATACCGTGCACTTTTGGACGCGAGGATCCAAAATGCTCAGAAATGTCAGGTGATG 4272  
Qy 4291 TGTCAACCCAGTGTCTCAAGTGTCTGCGAGCCTGCGGCTTGCCTGCTGAATATGTC 4350  
Db 4273 TGTCAACCCAGTGTCTCAAGTGTCTGCGAGCCTGCGGCTTGCCTGCTGAATATGTC 4332  
Qy 4351 ACACATTTACCGAGGCTTCTGCGTGACAAAATGAACCTCCAGGTCTCCAGACCAAG 4410  
Db 4333 ACACATTTACCGAGGCTTCTGCGTGACAAAATGAACCTCCAGGTCTCCAGACCAAG 4392  
Qy 4411 GAGCCAGCAGCAGCTTGCACTTGAAGGTCGATGAGTGGCCCGCAGGAATTAACAACGA 4470  
Db 4393 GAGCCAGCAGCAGCTTGCACTTGAAGGTCGATGAGTGGCCCGCAGGAATTAACAACGA 4452  
Qy 4471 GGACAGCAAGCTCGGACAGGAAGTACATTTCTCTGAGGAGTCAAAAAGTCTCATTTAT 4530  
Db 4453 GGACAGCAAGCTCGGACAGGAAGTACATTTCTCTGAGGAGTCAAAAAGTCTCATTTAT 4512  
Qy 4531 GACAAATGAAGCCAGAGAGCTGGAAGGCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4590  
Db 4513 GACAAATGAAGCCAGAGAGCTGGAAGGCGGTGGAAGAAATTTGAGCTGTGCTTCCC 4572  
Qy 4591 GACGGGATGATCTATTCTATGTCGCTTGGTCTTCCGAACTCGCAATACAGCCAAA 4650  
Db 4573 GACGGGATGATCTATTCTATGTCGCTTGGTCTTCCGAACTCGCAATACAGCCAAA 4632  
Qy 4651 GCA----- 4653  
Db 4633 GCAGATGTCCTATACATACTGAAGATGGAATCTCACCCGACACCACTGCTGTCGCCCGG 4692  
Qy 4654 ----- 4653  
Db 4693 AGAAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAAACAGCGCTGGGTACCGCCTTA 4752  
Qy 4654 -----GAAAAAGCAGAAAGCTGATGCTAAAACTG 4680  
Db 4753 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGAAAAAAGCAGAGCTGATGCTAACTG 4812

4681	CTTGGAAACTCCCTGCTGTAACCTGGAAAGGTGATGACCGTCTAGACATGAACTGCACGCTG	474
4813	CTTGGAAACTCCCTGCTGTAACCTGGAAAGGTGATGACCGTCTAGACATGAACTGCACGCTG	4872
4741	CCCTTCAGTGACCAGGTGGTGGTGGCGGCACCGAGGAAGGGCTCTACGCCCTGATGTC	4800
4873	CCCTTCAGTGACCAGGTGAGTGGTGGGCACCGAGGAAGGGCTCTACGCCCTGATGTC	4832
4801	TTGAAAAACTCCCTTAACCCCATGTCCCAGGAATTGGAGCAGTCTTCCAAATTTATATTATC	4860
4933	TTGAAAAACTCCCTTAACCCCATGTCCCAGGAATTGGAGCAGTCTTCCAAATTTATATTATC	4992
4861	AAGGACCTGGAGAGACTACTCATGATAGCAGGAGAGAGCGGGCAGCTGTCTTTGTGGAC	4820
4993	AAGGACCTGGAGAGACTACTCATGATAGCAGGTGAGAGCGGGCAGCTGTCTTTGTGGAC	5052
4921	GTGAAGAAAGTGAACACAGTCCCTGGCCAGTCCCACTGCTGCCACGCCGACATCTCA	4980
5053	GTGAAGAAAGTGAACACAGTCCCTGGCCAGTCCCACTGCTGCCACGCCGACATCTCA	5112
4981	CCCAACAATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTTGGGGAGCGCAAGATTGAGAAC	5040
5113	CCCAACAATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTTGGGGAGCGCAAGATTGAGAAC	5172
5041	GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTCTGCTCAATCTCGCTACACAGAAAC	5100
5173	GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTCTGCTCAATCTCGCTACACAGAAAC	5232
5101	CTCAGCAAAATCTGCATTCGGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5160
5233	CTCAGCAAAATCTGCATTCGGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5292
5161	TTCAACCAATTACAGTATCCCTCATTTGGAAACCAATAAATTTCTACCAAAATCGACATGAGCAG	5220
5293	TTCAACCAATTACAGTATCCCTCATTTGGAAACCAATAAATTTCTACCAAAATCGACATGAGCAG	5352
5221	TACACGCTCGAGGAATTCCTGATAGAAATGACCAATTCCTTTGGCACGTCGTGTGTTGCC	5280
5353	TACACGCTCGAGGAATTCCTGATAGAAATGACCAATTCCTTTGGCACGTCGTGTGTTGCC	5412
5281	GCCTCTTCGAACAGCTCCCTGTCTCAATCGTGCAGGTGNAACGCGCAGGGCAGCAGAG	5340
5413	GCCTCTTCGAACAGCTCCCTGTCTCAATCGTGCAGGTGNAACGCGCAGGGCAGCAGAG	5472
5341	GAGTACTTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAACAGCTAGC	5400
5473	GAGTACTTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAACAGCTAGC	5532
5401	CGCACAGACATCTCAAGTGGAGTCGTCTTACCTTTTGGCCCTTTGCCCTACAGAGAACCTTAT	5460
5533	CGCACAGACATCTCAAGTGGAGTCGTCTTACCTTTTGGCCCTTTGCCCTACAGAGAACCTTAT	5592
5461	CTGTTTGTGACCCACATTCAACTCACTCGAAGTAATTGAGATCCAGCAGCCTCCTCAGCA	5520
5593	CTGTTTGTGACCCACATTCAACTCACTCGAAGTAATTGAGATCCAGCAGCCTCCTCAGCA	5652
5521	GGGACCCCTGCCGAGCGTACCTGGACATCCCGAAACCCGGCTACCTGGGCCCTGCCATT	5580
5653	GGGACCCCTGCCGAGCGTACCTGGACATCCCGAAACCCGGCTACCTGGGCCCTGCCATT	5712
5581	TCCCTCAGGACGATTTACTTTGGCGTCTCATACCGAGTAAATTTAGGGTCAATTTCTCTCG	5640
5713	TCCCTCAGGACGATTTACTTTGGCGTCTCATACCGAGTAAATTTAGGGTCAATTTCTCTCG	5772
5641	AAGGGAACCTCGTGAAGGAGTCCGCACTGAACACCAACCGGGGCCCTGCCACCTCCCGC	5700
5773	AAGGGAACCTCGTGAAGGAGTCCGCACTGAACACCAACCGGGGCCCTGCCACCTCCCGC	5832
5701	AGCAGCCCAACAGCGAGGCCACCCAGTACACGACACATCACCAAGCGCTGGCC	5760
5833	AGCAGCCCAACAGCGAGGCCACCCAGTACACGACACATCACCAAGCGCTGGCC	5892
5761	TCCAGCCCAACGCGCCGCGCAAGGCCCCAGCCACCCCGCAGAGCGCAAGCACACCCACCGC	5820

[illegible]



Db 1378 TCTCAGGACAAAGTGTCTCAAGATGGAGCAGGAAATGACCCGGTTTACATCGGAGAGTGTCA 1437  
QY 1459 GAGGTGGAGGCTGTGCTTAGTTCAGAAAGGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1518  
Db 1438 GAGGTGGAGGCTGTGCTTAGTTCAGAAAGGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1497  
QY 1519 TCCCTCTGAGAGCAGGACCTTGCTTACTCATCAACAGAAATGCAAGTAAAGGCAAGT 1578  
Db 1498 TCCCTCTGAGAGCAGGACCTTGCTTACTCATCAACAGAAATGCAAGTAAAGGCAAGT 1557  
QY 1579 TTGAGCAAGCAGGATGGAGGTGTCCAGGAGGATGACAAAGCACTTGACGCTTCCAT 1638  
Db 1558 TTGAGCAAGCAGGATGGAGGTGTCCAGGAGGATGACAAAGCACTTGACGCTTCCAT 1617  
QY 1639 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAAGGCTCAA 1698  
Db 1618 GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAAGGCTCAA 1677  
QY 1699 GTGCAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1758  
Db 1678 GTGCAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1737  
QY 1759 CGGAGTGTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAAG 1818  
Db 1738 CGGAGTGTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAAG 1797  
QY 1819 CGGAAAGCGACAGATCTCAGCAATAAATCTTTGAAAGCTTAAGGATCAAGGAAAGCCTGAA 1878  
Db 1798 CGGAAAGCGACAGATCTCAGCAATAAATCTTTGAAAGCTTAAGGATCAGGGAAGCCTGAA 1857  
QY 1879 GTGCGAAATATGCGAAATCTGGAAGAGTCAATGCTGAGCAGAGCTCAAAATTTCAAGAG 1938  
Db 1858 GTGCGAAATATGCGAAATCTGGAAGAGTCAATGCTGAGCAGAGCTCAAAATTTCAAGAG 1917  
QY 1939 CTCGAGAGAAATCTGGAGAAAGGCT----- 1962  
Db 1918 CTCGAGAGAAATCTGGAGAAAGGCTGTAAAGCCAGACCGAGGCCACCGAGCTGTGTCAG 1977  
QY 1963 -----GCAAGAGCGAGCCGAGAGGAGCTGGAGAGCTGCAGAACCGAGAG 2010  
Db 1978 AATATCCGCGCAGGCAAGAGGCGAGCGAGGAGCTGGAGAGCTGCAGAACCGAGAG 2037  
QY 2011 GATTTCTTGAAGGCTCAGAAAGAGCTGGTGGAAAGCTGAGGAACCGCCGCTTCTCTG 2070  
Db 2038 GATTTCTTGAAGGCTCAGAAAGAGCTGGTGGAAAGCTGAGGAACCGCCGCTTCTCTG 2097  
QY 2071 GAGAAACAGGTAAGAGACTAGAGACGATGGAGCTGAGAGAAACAGACATGAAGATGAC 2130  
Db 2098 GAGAAACAGGTAAGAGACTAGAGACGATGGAGCTGAGAGAAACAGACATGAAGATGAC 2157  
QY 2131 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2190  
Db 2158 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAAGAG 2217  
QY 2191 AATCATCGGAGGCCCAAGTCTCAGCCAGACCTTAGAAGTGACCTGTAAAACAGAAAGAG 2250  
Db 2218 AATCATCGGAGGCCCAAGTCTCAGCCAGACCTTAGAAGTGACCTGTAAAACAGAAAGAG 2277  
QY 2251 CAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATTAAGAAAGACCTTGGCTGAC 2310  
Db 2278 CAGCACTATGAGGAAAGATTAAGTATTGGACATCAGATTAAGAAAGACCTTGGCTGAC 2337  
QY 2311 AAGGAGACATGAGAAACATGATGACAGACACGAGGAGGAGCCCATGAGAGGCGCAAA 2370  
Db 2338 AAGGAGACATGAGAAACATGATGACAGACACGAGGAGGAGGCCCATGAGAGGCGCAAA 2397  
QY 2371 ATTCTCAGGCAACAGAGGCGATGATCAATGCTATGCTATGCTTCAAGATCAGATCCCTGGAA 2430  
Db 2398 ATTCTCAGGCAACAGAGGCGATGATCAATGCTATGCTATGCTTCAAGATCAGATCCCTGGAA 2457  
QY 2431 CAGAGGATTTGGAAGTGTCTGAGGCAATTAACCTTTGACGCAATAGCAGTCTTTTACC 2490  
Db 2458 CAGAGGATTTGGAAGTGTCTGAGGCAATTAACCTTTGACGCAATAGCAGTCTTTTACC 2517

QY 2491 CAAAGGAAACATGAAGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2550  
Db 2518 CAAAGGAAACATGAAGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2577  
QY 2551 CTGGAGACACAGGCTGGGAAAGTTGGAGCCCGAGAACCGAAATCTGGAGGACAGCTGGAG 2610  
Db 2578 CTGGAGACACAGGCTGGGAAAGTTGGAGCCCGAGAACCGAAATCTGGAGGACAGCTGGAG 2637  
QY 2611 AAGATCACCCACCAAGACCAAGTCAAGAAATCGGCTGTGGAATCTGGAGACAAAGATTG 2670  
Db 2638 AAGATCACCCACCAAGACCAAGTCAAGAAATCGGCTGTGGAATCTGGAGACAAAGATTG 2697  
QY 2671 CGGAGGTCAGTCTAGAGCACGAGGAGCAGAAATCTGGAGCTCAAGCGCCAGCTCACAGAG 2730  
Db 2698 CGGAGGTCAGTCTAGAGCACGAGGAGCAGAAATCTGGAGCTCAAGCGCCAGCTCACAGAG 2757  
QY 2731 CTACAGCTCTCCCTGCGAGGCGGAGTCAAGTTGACGCCCTGACGGCTGCAAGCGGCG 2790  
Db 2758 CTACAGCTCTCCCTGCGAGGCGGAGTCAAGTTGACGCCCTGACGGCTGCAAGCGGCG 2817  
QY 2791 GCGCTGGAGAGCCAGCTTCCGAGGCGAAGACAGAGCTTGAAGAGACCAACAGCAGAAAGCT 2850  
Db 2818 GCGCTGGAGAGCCAGCTTCCGAGGCGAAGACAGAGCTTGAAGAGACCAACAGCAGAAAGCT 2877  
QY 2851 GAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAATTCAGCGCAATTTGATGCT 2910  
Db 2878 GAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAATTCAGCGCAATTTGATGCT 2937  
QY 2911 CTTCTGAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAACAGCAGCTGACCAG 2970  
Db 2938 CTTCTGAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAACAGCAGCTGACCAG 2997  
QY 2971 GACAAACGCTGAATCAAACCAAAATCTTCTACTGTGTCAAAACAACTCGATGAGGCTTCT 3030  
Db 2998 GACAAACGCTGAATCAAACCAAAATCTTCTACTGTGTCAAAACAACTCGATGAGGCTTCT 3057  
QY 3031 GCGCCCAACAGCAGATTTGACAACTCGAAGTGAAGTGAACCATCTCCGCGGAGATC 3090  
Db 3058 GCGCCCAACAGCAGATTTGACAACTCGAAGTGAAGTGAACCATCTCCGCGGAGATC 3117  
QY 3091 ACGAAACAGAGATGCGAGCTTACAGCCAGAGCAAAACGATGGAGCTCTGAAGACACG 3150  
Db 3118 ACGAAACAGAGATGCGAGCTTACAGCCAGAGCAAAACGATGGAGCTCTGAAGACACG 3177  
QY 3151 TGCACCATGCTGGAGAACAGGTCATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3210  
Db 3178 TGCACCATGCTGGAGAACAGGTCATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3237  
QY 3211 AAAGAGCGGAGTGGGAGGCTGGAGGAGCTCTCGGCTGATGAGAAATCCGAGTTTGA 3270  
Db 3238 AAAGAGCGGAGTGGGAGGCTGGAGGAGCTCTCGGCTGATGAGAAATCCGAGTTTGA 3297  
QY 3271 TGTGCGGTTGAGAGCTGAGAGATGCTGACACCGAGAAACAGAGCAGGCGGAGAGCC 3330  
Db 3298 TGTGCGGTTGAGAGCTGAGAGATGCTGACACCGAGAAACAGAGCAGGCGGAGAGCC 3357  
QY 3331 GATCAGCGGATCACCGAGTCTCGCGAGGCTGGAGGAGCTCTCGGCTGATGAGAAATCCGAGTTTGA 3390  
Db 3358 GATCAGCGGATCACCGAGTCTCGCGAGGCTGGAGGAGCTCTCGGCTGATGAGAAATCCGAGTTTGA 3417  
QY 3391 GAGATTCGCTGCGAGGCTCTCAAGAGCAGAGCTGAAGCGCGAGAGCTCTCT 3450  
Db 3418 GAGATTCGCTGCGAGGCTCTCAAGAGCAGAGCTCTCAAGAGCAGAGCTCTCT 3477  
QY 3451 GACAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAATGCTCCGAGCTTA 3510  
Db 3478 GACAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAATGCTCCGAGCTTA 3537  
QY 3511 CAGCAGAGCTGAGAGCTGAACAGAGCTCAACAGAGGCTTCTGGAAGAGCAGGCAAA 3570  
Db 3538 CAGCAGAGCTGAGAGCTGAACAGAGCTCAACAGAGGCTTCTGGAAGAGCAGGCAAA 3597

QY 3571 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCCTGCTGACTCAAGGACTGCAG 3630  
DB 3598 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCCTGCTGACTCAAGGACTGCAG 3657  
QY 3631 GAAGCTCTAGATCGGGCTGATCTAATGAAGACAGAAAAAGTGAATTTGAGTATCAGCTG 3690  
DB 3658 GAAGCTCTAGATCGGGCTGATCTAATGAAGACAGAAAAAGTGAATTTGAGTATCAGCTG 3717  
QY 3691 GAAAAATTCAGGTTCTCTATTCTCATGAAAAAGTGAATTTGAGGAGCACTATTCTCAA 3750  
DB 3718 GAAAAATTCAGGTTCTCTATTCTCATGAAAAAGTGAATTTGAGGAGCACTATTCTCAA 3777  
QY 3751 CAAACCAAACTCATTTGATTTCTGCAAGCCAAAATGGACCAACCTGCTAAAAAGAAAAAG 3810  
DB 3778 CAAACCAAACTCATTTGATTTCTGCAAGCCAAAATGGACCAACCTGCTAAAAAGAAAA-- 3835  
QY 3811 GGTATTATTAGTCGAGGAAAGGACCTGCTTTACCCACACAGAGTTCCTCTGAGTAC 3870  
DB 3836 -----AGTGGCTCTGAGTAC 3852  
QY 3871 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTAGAGAAAGCC 3930  
DB 3853 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTAGAGAAAGCC 3912  
QY 3931 CTTCAGAAACCCGATCAGCTCCGGTCCGCCCGGAGGAGCTGCCACCCGAAAGCA 3990  
DB 3913 CTTCAGAAACCCGATCAGCTCCGGTCCGCCCGGAGGAAAGCTGCCACCCGAAAGCA 3972  
QY 3991 ACGAGCCACCCACACCCATCCAGCCAGCCACCGGAGGACAGATGCCATGTCGCC 4050  
DB 3973 ACGAGCCACCCACACCCATCCAGCCAGCCACCGGAGGACAGATGCCATGTCGCC 4032  
QY 4051 ATGTCGGTGCAGAGACCAAGCCAGTGCATGAGCTCTGCGCCCGCCATCCAGC 4110  
DB 4033 ATGTCGGTGCAGAGACCAAGCCAGTGCATGAGCTCTGCGCCCGCCATCCAGC 4092  
QY 4111 CGCAGAAAGGATCTCAATCCAGAGGAATTTAGTCGGCGCTTAAGGAACGCATGCAC 4170  
DB 4093 CGCAGAAAGGATCTCAATCCAGAGGAATTTAGTCGGCGCTTAAGGAACGCATGCAC 4152  
QY 4171 CACAATATTCTCACCAGATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTCTGTG 4230  
DB 4153 CACAATATTCTCACCAGATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTCTGTG 4212  
QY 4231 TGTCTGGATACCGTGACCTTTGAGACCCAGGATCCAAATGTCGAAATGTCAGGTGATG 4290  
DB 4213 TGTCTGGATACCGTGACCTTTGAGACCCAGGATCCAAATGTCGAAATGTCAGGTGATG 4272  
QY 4291 TGTACCCCAAGTGTCCAGCTGCTTGCAGCCACCTGCGGCTTGCTGCTGAATATGCC 4350  
DB 4273 TGTACCCCAAGTGTCCAGCTGCTTGCAGCCACCTGCGGCTTGCTGCTGAATATGCC 4332  
QY 4351 ACACATTCACCGAGGCTTCTGCCGTGACAAAATGAATCCCCAGGTCCTCAGACCCAG 4410  
DB 4333 ACACATTCACCGAGGCTTCTGCCGTGACAAAATGAATCCCCAGGTCCTCAGACCCAG 4392  
QY 4411 GAGCCAGAGCAGCTTGACCTGGAAGGTTGATGAAGTGCACAGGATTAACAAACGA 4470  
DB 4393 GAGCCAGAGCAGCTTGACCTGGAAGGTTGATGAAGTGCACAGGATTAACAAACGA 4452  
QY 4471 GGAACCAAGGCTGGACAGGAAGTACATTTGCTCTGGAGGATCAAAAGTCTCATTTAT 4530  
DB 4453 GGAACCAAGGCTGGACAGGAAGTACATTTGCTCTGGAGGATCAAAAGTCTCATTTAT 4512  
QY 4531 GACAATGAAGCCAGAGAGCTGACAGAGCCGGTGGAGAAATTTGAGCTGCTCTGCC 4590  
DB 4513 GACAATGAAGCCAGAGAGCTGACAGAGCCGGTGGAGAAATTTGAGCTGCTCTGCC 4572  
QY 4591 GACGGGATGTATCTATTATGATGCGCTGCTGCTTCCGAAGTCCGAATACAGCCAAA 4650  
DB 4573 GACGGGATGTATCTATTATGATGCGCTGCTGCTTCCGAAGTCCGAATACAGCCAAA 4632  
QY 4651 GCA----- 4653

DB 4633 GCAGATGTCCCATACATACTGAAGATGGAATCTCACCCGCACACCACTGCTGCGCCGG 4692  
QY 4654 ----- 4653  
DB 4693 AGAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGTCAACCGCTTA 4752  
QY 4654 -----GAAAAGCAGAACTGATGCTAAACTG 4680  
DB 4753 GAATCAGTTGTCGCGAGGTGGAGAGTTTCTAGGAAAAAAGCAGAACTGATGCTAAACTG 4812  
QY 4681 CTTGGAAATCTCCCTGCTGAAACTGGAAGGTGATGACCGCTTAGACATGAATGCAACCGCTG 4740  
DB 4813 CTTGGAAATCTCCCTGCTGAAACTGGAAGGTGATGACCGCTTAGACATGAATGCAACCGCTG 4872  
QY 4741 CCCTTCAGTCACACAGGTGCTGTTGGTGGGCACCGAGAAAGGCTCTACGCGCTGAATGTC 4800  
DB 4873 CCCTTCAGTCACACAGGTGCTGTTGGTGGGCACCGAGAAAGGCTCTACGCGCTGAATGTC 4932  
QY 4801 TTGAAAACTCCCTAACCCATGTCCTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4860  
DB 4933 TTGAAAACTCCCTAACCCATGTCCTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4992  
QY 4861 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGCGGACACTGTGCTTGTGGAC 4920  
DB 4993 AAGGACCTGGAGAGCTACTCATGATAGCAGGTGAAGAGCGGACACTGTGCTTGTGGAC 5052  
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 4980  
DB 5053 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 5112  
QY 4981 CCCAACATTTTGAAGTGTCAAGGGCTGCCACTGTTTGGGCGAGCAAGATTGAGAAC 5040  
DB 5113 CCCAACATTTTGAAGTGTCAAGGGCTGCCACTGTTTGGGCGAGCAAGATTGAGAAC 5172  
QY 5041 GGGCTCTGCATCTGTGAGCCATCCCGAGCAAAAGTCTCATTTCTCCGCTACAAAGAAAAC 5100  
DB 5173 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTCTCATTTCTCCGCTACAAAGAAAAC 5232  
QY 5101 CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5160  
DB 5233 CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5292  
QY 5161 TTCACCAATTTACAGTATCTCATTTGGAACCAATTAATTTCTACGAAATCGACATGAAGCAG 5220  
DB 5293 TTCACCAATTTACAGTATCTCATTTGGAACCAATTAATTTCTACGAAATCGACATGAAGCAG 5352  
QY 5221 TACAGCTCGAGGAATTCCTGGATGAAGATGAACATTCCTTTGGCACTGTGTTTGGCC 5280  
DB 5353 TACAGCTCGAGGAATTCCTGGATGAAGATGAACATTCCTTTGGCACTGTGTTTGGCC 5412  
QY 5281 GCCTCTTCCACAGCTTCCCTGCTCAATCGTCAGGTGAACCGCAGCGCAGCGAGAG 5340  
DB 5413 GCCTCTTCCACAGCTTCCCTGCTCAATCGTCAGGTGAACCGCAGCGCAGCGAGAG 5472  
QY 5341 GAGTACTTGTGTTTCCAGGAATTTGGAGTGTTCGTGGATTTCTACGGAAGACGTAGC 5400  
DB 5473 GAGTACTTGTGTTTCCAGGAATTTGGAGTGTTCGTGGATTTCTACGGAAGACGTAGC 5532  
QY 5401 CGCAGACAGATCTCAAGTGGAGTTCGCTTACCTTTGGCCCTTGCCTACAGAAACCTTAT 5460  
DB 5533 CGCAGACAGATCTCAAGTGGAGTTCGCTTACCTTTGGCCCTTGCCTACAGAAACCTTAT 5592  
QY 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATGAGATCCAGGACCGCTCTCAGCA 5520  
DB 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATGAGATCCAGGACCGCTCTCAGCA 5652  
QY 5521 GGGACCCCTGCGCAGCGTACTCGAATCCCGAAACCCGCGCTACCTGGGCGCTGCCATT 5580  
DB 5653 GGGACCCCTGCGCAGCGTACTCGAATCCCGAAACCCGCGCTACCTGGGCGCTGCCATT 5712  
QY 5581 TCCTCAGAGAGCAATTTACTTGGCGTCTCATACAGGATAAATTAAGGTCATTTGCTGC 5640

Db 5713 TCCTCAGGAGGATTACTTGGGGTCTCTATACAGGATAAATTAAGGTCATTGCTGC 5772  
Qy 5641 AAGGAAACCTCTGTAAGGAGTCGGCACTGAACACACCGGGGCCCGTCCACTCCCGC 5700  
Db 5773 AAGGAAACCTCTGTAAGGAGTCGGCACTGAACACACCGGGGCCCGTCCACTCCCGC 5832  
Qy 5701 AGCAGCCCCACAGAGGAGGCGCCACACCTACACGAGCAGATCACCAGCGCTGGCC 5760  
Db 5833 AGCAGCCCCACAGAGGAGGCGCCACACCTACACGAGCAGATCACCAGCGCTGGCC 5892  
Qy 5761 TCCAGCCAGCGCCCGCCAGAGGCCCGCCAGCCACCGCGAGAGGCAACACCCACCGC 5820  
Db 5893 TCCAGCCAGCGCCCGCCAGAGGCCCGCCAGCCACCGCGAGAGGCAACACCCACCGC 5952  
Qy 5821 TACCGGAGGGGCGGAGCGAGCTGCGCAGGAGCAAGTCTCTGGCGCGCCCGCTGGAGCGA 5880  
Db 5953 TACCGGAGGGGCGGAGCGAGCTGCGCAGGAGCAAGTCTCTGGCGCGCCCGCTGGAGCGA 6012  
Qy 5881 GAGAAGTCCCGCGCGGATGCTCAGACACCGGAGAGCGGTCCTCCCGGAGGCTGTTT 5940  
Db 6013 GAGAAGTCCCGCGCGGATGCTCAGACACCGGAGAGCGGTCCTCCCGGAGGCTGTTT 6072  
Qy 5941 GAAGACAGCAGCGCGCGGCTGCTGCGGAGCGCGTGTAGGACCGCGTGTCCAGGTT 6000  
Db 6073 GAAGACAGCAGCGCGCGGCTGCTGCGGAGCGCGTGTAGGACCGCGTGTCCAGGTT 6132  
Qy 6001 AACAGGGAAGAGGCA 6017  
Db 6133 AACAGGTGAGGAGCA 6149

RESULT 11  
ADA05653  
ID ADA05653 standard; cDNA; 6201 BP.  
XX  
AC ADA05653;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human NOV1g encoding cDNA SEQ ID NO:13.  
XX  
KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipaeamic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..6201  
FT /tag= a  
FT /product= "NOV1g"  
XX  
PN WO2003029424-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031373.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328949P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
XX

PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
PR 01-OCT-2002; 2002US-00262511.  
XX  
XX (CUPA-) CURAGEN CORP.  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dippio VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2003-381626/36.  
DR P-PSDB; ADA05654.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
PS Claim 20; Page 104-105; 586pp; English.  
XX  
CC The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell  
CC comprising the above vector; (6) an antibody that immunospecifically  
CC binds to the polypeptide described above; (7) methods for determining the  
CC presence or amount of the above polypeptide or nucleic acid molecule in a  
CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above  
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
CC method of identifying an agent that binds to the polypeptide described  
CC above; (10) a method for identifying a potential therapeutic agent for  
CC use in treating a pathology that is related to an aberrant expression or  
CC aberrant physiological interactions of the polypeptide; (11) a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide; (12) a method for modulating  
CC the activity of the polypeptide described above; (13) methods of treating  
CC or preventing a pathology associated with the above polypeptide in a  
CC mammal; and (14) a method for producing the above polypeptide. NOVX  
CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian  
CC and antilipaeamic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence encodes a human NOVX protein from  
CC the present invention.  
XX

SQ	Sequence	6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;
	Query Match	85.6%; Score 5629.8; DB 7; Length 6201;
	Best Local Similarity	95.7%; Pred. No. 0;
	Matches 5929; Conservative	0; Mismatches 22; Indels 246; Gaps 4;
QY	19	ATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGCTGAACCCATT 78
DB	1	ATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGCTGAACCCATT 60
QY	79	GCACGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTTATGACTCAA 138
DB	61	GCACGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTTATGACTCAA 120
QY	139	CAGCAGATGTCCTCTCTTCCCGAGAAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198
DB	121	CAGCAGATGTCCTCTCTTCCCGAGAAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
QY	199	GAATGAGTCAGCTGCTCTGATGAAGATTAAGCACTGAGCAACTTTGTCCGGAGTAT 258
DB	181	GAATGAGTCAGCTGCTCTGATGAAGATTAAGCACTGAGCAACTTTGTCCGGAGTAT 240
QY	259	TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 318
DB	241	TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 300
QY	319	AGTCTTGTAGTTGTGTCACCTTTTCTGAAAGTGCAGGTGGTAAGAGAGAAAACCGGG 378
DB	301	AGTCTTGTAGTTGTGTCACCTTTTCTGAAAGTGCAGGTGGTAAGAGAGAAAACCGGG 360
QY	379	GACATCTAGCTATGAAAGTGTGAAGAGAGGCTTTATTTGGCCCGAGGACAGGTTTCA 438
DB	361	GACATCTAGCTATGAAAGTGTGAAGAGAGGCTTTATTTGGCCCGAGGACAGGTTTCA 420
QY	439	TTTTTTGAGGAAGCGGAACATATTATCTCGAAGCACAAGCCCGCTGGATCCCCAATTA 498
DB	421	TTTTTTGAGGAAGCGGAACATATTATCTCGAAGCACAAGCCCGCTGGATCCCCAATTA 480
QY	499	CAGTATGCTTTACAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 558
DB	481	CAGTATGCTTTACAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540
QY	559	GACTTGCTCTACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAACCTGATACAGTTT 618
DB	541	GACTTGCTCTACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAACCTGATACAGTTT 600
QY	619	TACCTAGCTGAGCTGATTTTGGCTCTTACACGCTTCATCTGATGGGATAGTGATCGA 678
DB	601	TACCTAGCTGAGCTGATTTTGGCTCTTACACGCTTCATCTGATGGGATAGTGATCGA 660
QY	679	GACATCAAGCCTGAGAACATTTCTGTTGACCGCAGGACACATCAAGCTGGTGGATTTT 738
DB	661	GACATCAAGCCTGAGAACATTTCTGTTGACCGCAGGACACATCAAGCTGGTGGATTTT 720
QY	739	GGATCTGCCCGGAAAATGAATTCAAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 798
DB	721	GGATCTGCCCGGAAAATGAATTCAAAACA--GGTGAATGCCAAACTCCCGATTGGGACC 777
QY	799	CAGATTAATAGGCTCCTGAAGTCTGATGTAAGACGGGATGGAAGGACCTAC 858
DB	778	CCAGATTAATAGGCTCCTGAAGTCTGATGTAAGACGGGATGGAAGGACCTAC 837
QY	859	GGCTCGGACTGTGACTGGTGGTCAGTGGGCTGATTGCCCTATGAGATGATTTATGGGAGA 918
DB	838	GGCTCGGACTGTGACTGGTGGTCAGTGGGCTGATTGCCCTATGAGATGATTTATGGGAGA 897
QY	919	TCCCCCTTCGAGAGGAACTCTGCGCAAGACCTTCAATTAACATTAATGATTTTCAGCGG 978
DB	898	TCCCCCTTCGAGAGGAACTCTGCGCAAGACCTTCAATTAACATTAATGATTTTCAGCGG 957
QY	979	TTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGTATCTGATTCATAAGC 1038
DB	958	TTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGTATCTGATTCATAAGC 1017

QY	1039	TTGTTGTGCGGCAGAAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC 1098
DB	1018	TTGTTGTGCGGCAGAAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCCTTTCTTC 1077
QY	1099	TCTAAATTTGACTGGAACAACATTTGTAATCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1158
DB	1078	TCTAAATTTGACTGGAACAACATTTGTAACGCTCTCTCCCTTCGTTCCCAACCTCAAG 1137
QY	1159	TCTGACGATCAGACCTCCAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTCATCTCT 1218
DB	1138	TCTGACGATCAGACCTCCAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTCATCTCT 1197
QY	1219	CGGTGCGAGCTGAGCCCTCAGGCTTCTCGGTGAAAGAACTGCGTTTGTGGGGTTTTTCG 1278
DB	1198	CGGTGCGAGCTGAGCCCTCAGGCTTCTCGGTGAAAGAACTGCGTTTGTGGGGTTTTTCG 1257
QY	1279	TACAGCAAGCACTGGGATTTCTTGTAGATCTGAGTCTGTTCTGTCGGCTCTGGAATCC 1338
DB	1258	TACAGCAAGCACTGGGATTTCTTGTAGATCTGAGTCTGTTCTGTCGGCTCTGGAATCC 1317
QY	1339	CCTGCCAAGACTAGCTCCATCGAAAGAAACTTCTCATCAAAAAGCAAGAGCTACAAGAC 1398
DB	1318	CCTGCCAAGACTAGCTCCATCGAAAGAAACTTCTCATCAAAAAGCAAGAGCTACAAGAC 1377
QY	1399	TCTCAGACAAGTGTCAACAGATGGAGCAAGAAATGACCCGGTTACATCGGAGAGTGCA 1458
DB	1378	TCTCAGACAAGTGTCAACAGATGGAGCAAGAAATGACCCGGTTACATCGGAGAGTGCA 1437
QY	1459	GAGGTGAGGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGGCTCTGAGAGCTCAGAGA 1518
DB	1438	GAGGTGAGGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGGCTCTGAGAGCTCAGAGA 1497
QY	1519	TCCCTCTCGAGCAGGACCTTGTCTACATACAGAAATGAGAGTCTTAAAGCGAAGT 1578
DB	1498	TCCCTCTCGAGCAGGACCTTGTCTACATACAGAAATGAGAGTCTTAAAGCGAAGT 1557
QY	1579	TTGAGCAAGCAGGATGGAGGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCCAT 1638
DB	1558	TTGAGCAAGCAGGATGGAGGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCCAT 1617
QY	1639	GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGGCTCAA 1698
DB	1618	GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGGCTCAA 1677
QY	1699	GTGGAAGAAATGAGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1758
DB	1678	GTGGAAGAAATGAGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1737
QY	1759	CGGAGTCTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAAATTCAG 1818
DB	1738	CGGAGTCTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAAATTCAG 1797
QY	1819	CGGAAAGCGACAGAAATCTCAGATTAACCTGTTGAAGCTAAGGATCAAGGGAAGCTGAA 1878
DB	1798	CGGAAAGCGACAGAAATCTCAGATTAACCTGTTGAAGCTAAGGATCAAGGGAAGCTGAA 1857
QY	1879	GTGGGAAATATGCGAACTGGGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1938
DB	1858	GTGGGAAATATGCGAACTGGGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1917
QY	1939	CTCCAAGAGAACTGGGAAGGCT----- 1962
DB	1918	CTCCAAGAGAACTGGGAAGGCTGTAAAAGCCAGCAGCGGCCCCAGCTGCTCAG 1977
QY	1963	-----GCAAGAGGAGCGAGCGAGGGAGCTGAGAGCTGAGAGCTGCAGAACCGAGAG 2010
DB	1978	AATATCCCGCAGGCAAGGAGCGAGCGAGGGAGCTGGAGAAGCTGCAGAACCGAGAG 2037
QY	2011	GATTCTTTCTAAGGCATCAGAAAAGAGCTGGTGGAGCTGAGGAACGCCCATTTCTG 2070
DB	2038	GATTCTTTCTAAGGCATCAGAAAAGAGCTGGTGGAGCTGAGGAACGCCCATTTCTG 2097

2071 GAGAACAAAGGTAAGAGACTAGAGACCATGGAGCGTAGAGAAAAACAGACTGAAGGATGAC 2130  
Db  
2098 GAGAACAAAGGTAAGAGACTAGAGACCATGGAGCGTAGAGAAAAACAGACTGAAGGATGAC 2157  
Qy  
2131 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAGAG 2190  
Db  
2158 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGGAGCTCGAGAG 2217  
Qy  
2191 AAACATCGGAGGCCCCAAGTCTCAGCCCCAGCACCTAGAAAGTGCACCTGAAACAGAAAGAG 2250  
Db  
2218 AAACATCGGAGGCCCCAAGTCTCAGCCCCAGCACCTAGAAAGTGCACCTGAAACAGAAAGAG 2277  
Qy  
2251 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATTAAGAAAGACCTGGCTGAC 2310  
Db  
2278 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATTAAGAAAGACCTGGCTGAC 2337  
Qy  
2311 AAGGAGACACTGAGAAACATGATGCAGACACAGAGAGGAGGCCCATGAGAAAGGCCAAA 2370  
Db  
2338 AAGGAGACACTGAGAAACATGATGCAGACACAGAGAGGAGGCCCATGAGAAAGGCCAAA 2397  
Qy  
2371 ATTCTAGCGAAACAGAAAGCGATGATCAATGCTATGATTCAGATTCAGATCCCTGGAA 2430  
Db  
2398 ATTCTAGCGAAACAGAAAGCGGATGATCAATGCTATGATTCAGATTCAGATCCCTGGAA 2457  
Qy  
2431 CAGAGGATTCGGAACCTCTGAAGCCCAATAGACTTGCAGCAATAGCAGTCTTTTACC 2490  
Db  
2458 CAGAGGATTCGGAACCTCTGAAGCCCAATAGACTTGCAGCAATAGCAGTCTTTTACC 2517  
Qy  
2491 CAAGAAGAACATGAAGCCCCAAGAGAGATGATTTCTGAATTCAGGCAACAGAAATTTTAC 2550  
Db  
2518 CAAGAAGAACATGAAGCCCCAAGAGAGATGATTTCTGAATTCAGGCAACAGAAATTTTAC 2577  
Qy  
2551 CTGAGACACAGCTGGAGATTTGGAGGCCCAAGAACCGGAATCTGGAGAGCAGCTGGAG 2610  
Db  
2578 CTGAGACACAGCTGGAGATTTGGAGGCCCAAGAACCGGAATCTGGAGAGCAGCTGGAG 2637  
Qy  
2611 AAGATCAGGCCAACAGACACAGTGAACAAGATTCGGCTCTGGAATCGAGACCAAGATTG 2670  
Db  
2638 AAGATCAGGCCAACAGACACAGTGAACAAGATTCGGCTCTGGAATCGAGACCAAGATTG 2697  
Qy  
2671 CGGAGGTCAGTCTAGAGACAGAGAGCAAGAACTGGAGCTCAAGCCGACGCTCAGAG 2730  
Db  
2698 CGGAGGTCAGTCTAGAGACAGAGAGCAAGAACTGGAGCTCAAGCCGACGCTCAGAG 2757  
Qy  
2731 CTACAGCTCTCCCTCAGAGAGCGAGTCAAGTTCAGAGCCCTGCAGGCTCAGCGGCG 2790  
Db  
2758 CTACAGCTCTCCCTCAGAGAGCGAGTCAAGTTCAGAGCCCTGCAGGCTCAGCGGCG 2817  
Qy  
2791 GCCTCGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGACCAACAGCAAGCT 2850  
Db  
2818 GCCTCGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGACCAACAGCAAGCT 2877  
Qy  
2851 GAAGAGAGATCAGGACCTCAGGACATAGAGATGAATCCAGGCAAAATTCATGCT 2910  
Db  
2878 GAAGAGAGATCAGGACCTCAGGACATAGAGATGAATCCAGGCAAAATTCATGCT 2937  
Qy  
2911 CTTTCGTAACAGCTGTACTGTAACTCAGACACCTGGAGGACGCTAAACAGCTGACCGAG 2970  
Db  
2938 CTTTCGTAACAGCTGTACTGTAACTCAGACACCTGGAGGACGCTAAACAGCTGACCGAG 2997  
Qy  
2971 GACAACTGTAACCAACCAAACTTCTACTTGTCCAAACCACTCGATGAGCTTCT 3030  
Db  
2998 GACAACTGTAACCAACCAAACTTCTACTTGTCCAAACCACTCGATGAGCTTCT 3057  
Qy  
3031 GGGCCCAACGACAGATTTGTAACCTGCGAAGTGAAGTGGACCACTCTCCGCCGGAGATC 3090  
Db  
3058 GGGCCCAACGACAGATTTGTAACCTGCGAAGTGAAGTGGACCACTCTCCGCCGGAGATC 3117  
Qy  
3091 ACGGAACGAGAGATGAGCTTACCAGCCAGAGCAAAACGATGGAGGCTCTGAAAGCCAG 3150  
Db  
3118 ACGGAACGAGAGATGAGCTTACCAGCCAGAGCAAAACGATGGAGGCTCTGAAAGCCAG 3177  
Qy  
3151 TGCACCATGCTGGAGAACAGGTCATGGATTTGGAGGCCCTTAACCAATGATGCTGTAGAA 3210

3178 TGCCACATGCTGGAGAACAGGTCATGGATTTGGAGGCCCTTAAACAGATGAGCTGTAGAA 3237  
Qy  
3211 AAAGAGCGCAGTGGGAGGCTCGAGGAGCGTCTGGGTGATGAGAAATCCAGTTTGTAG 3270  
Db  
3238 AAAGAGCGCAGTGGGAGGCTCGAGGAGCGTCTGGGTGATGAGAAATCCAGTTTGTAG 3297  
Qy  
3271 TGTGGGTTCCAGAGCTGCAGAGAAATCTGGACACCGAGAAAACAGACAGCGGCGAGAGCC 3330  
Db  
3298 TGTGGGTTCCAGAGCTGCAGAGAAATCTGGACACCGAGAAAACAGACAGCGGCGAGAGCC 3357  
Qy  
3331 CATCAGCGGATCACCGAGTCTCCGAGGTGGAGCTGGAGCTGGAGGAGCAAGGCT 3390  
Db  
3358 GATCAGCGGATCACCGAGTCTCCGAGGTGGAGCTGGAGCTGGAGGAGCAAGGCT 3417  
Qy  
3391 GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3450  
Db  
3418 GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3477  
Qy  
3451 GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAGGCTTA 3510  
Db  
3478 GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAGGCTTA 3537  
Qy  
3511 CAGCAGAGCTGGAGACTGAACAGAGAGCTCAAACAGAGGCTTCTGGAAAGAGCAAGCCAAA 3570  
Db  
3538 CAGCAGAGCTGGAGACTGAACAGAGAGCTCAAACAGAGGCTTCTGGAAAGAGCAAGCCAAA 3597  
Qy  
3571 TTAACAGCAGAGCTGGAGCTGCAGAAAAATCAATTTTCGCTGACTCAAGGACTGCAA 3630  
Db  
3598 TTAACAGCAGAGCTGGAGCTGCAGAAAAATCAATTTTCGCTGACTCAAGGACTGCAA 3657  
Qy  
3631 GAAGCTCTAGATCGGCTGATCTACTGAACAGAAAGAGTACCTTGGAGTATCAGCTG 3690  
Db  
3658 GAAGCTCTAGATCGGCTGATCTACTGAACAGAAAGAGTACCTTGGAGTATCAGCTG 3717  
Qy  
3691 GAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAAAAAGGAGCACTATTCTCAA 3750  
Db  
3718 GAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAAAAAGGAGCACTATTCTCAA 3777  
Qy  
3751 CAACACAAAATCTAGTTGATTTCTGCAGCCAAAATGCAACCTCTTAAAGAAAG 3810  
Db  
3778 CAACACAAAATCTAGTTGATTTCTGCAGCCAAAATGCAACCTCTTAAAGAAAG -- 3835  
Qy  
3811 GGTTTATTTAGTCGACGGAAGAGGACCTGCTTTTACCCACAGAGTTCTCTGCAAGTAC 3870  
Db  
3836 -----AGGTGCTCTGCAAGTAC 3852  
Qy  
3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTGCTGTGCAGAGCTAGAGGAGCC 3930  
Db  
3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTGCTGTGCAGAGCTAGAGGAGCC 3912  
Qy  
3931 CTTCAGAAAGCCCGCATCGAGCTCCGCTCCCGGAGGAAGCTGCCCAACCAAGCA 3990  
Db  
3913 CTTCAGAAAGCCCGCATCGAGCTCCGCTCCCGGAGGAAGCTGCCCAACCAAGCA 3972  
Qy  
3991 ACGSACCAACCAACCACTCCAGCCAGCCACCGGAGGAGCAGATTCGCAATGTCGCC 4050  
Db  
3973 ACGSACCAACCAACCACTCCAGCCAGCCACCGGAGGAGCAGATTCGCAATGTCGCC 4032  
Qy  
4051 ATCGTGGGTCGCGAGAGCAGCCAGGCTGAGCTGAGCTGCTGCCCGCCCATCCAGC 4110  
Db  
4033 ATCGTGGGTCGCGAGAGCAGCCAGGCTGAGCTGAGCTGCTGCCCGCCCATCCAGC 4092  
Qy  
4111 CGCAGAAAGGAGTCTTCACTCCAGAGGAAATTTAGTCGGCGTCTTAAAGGAAACGATGCAC 4170  
Db  
4093 CGCAGAAAGGAGTCTTCACTCCAGAGGAAATTTAGTCGGCGTCTTAAAGGAAACGATGCAC 4152  
Qy  
4171 CACAAATTTCTCAGGATTTCAAGCTAGGACTGAAATCGGAGCCCAAGAGTGTGCTG 4230  
Db  
4153 CACAAATTTCTCAGGATTTCAAGCTAGGACTGAAATCGGAGCCCAAGAGTGTGCTG 4212  
Qy  
4231 TGTCTGATACCGTGCACCTTTGACGCCAGGCACTCCAAATGCTCGAATGTCTCAGGATG 4290

Db 4213 TGTCTGGATACCGTGACACTTTGGACGCCAGGCATCCAAATGTCTAGAAATGTCAGGTGATG 4272  
Qy 4291 TGTACCCCAAGTGTCTCACGCTGTTGCCAGCCACTCGCGCTTGCCTGTGCTGAATATGCC 4350  
Db 4273 TGTACCCCAAGTGTCTCACGCTGTTGCCAGCCACTCGCGCTTGCCTGTGCTGAATATGCC 4332  
Qy 4351 ACACACTTCAACCGAGGCTTCTGCCGTGACAAATGAACTCCCCACAGGTCTCCAGACCAAG 4410  
Db 4333 ACACACTTCAACCGAGGCTTCTGCCGTGACAAATGAACTCCCCACAGGTCTCCAGACCAAG 4392  
Qy 4411 GAGCCACAGCAGCTTGCACCTGGAAGGTTGATGAAGGTGCCCCAGGAATAACAAACGA 4470  
Db 4393 GAGCCACAGCAGCTTGCACCTGGAAGGTTGATGAAGGTGCCCCAGGAATAACAAACGA 4452  
Qy 4471 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTTGGAGGATCAAAAGTCTCTCATTTAT 4530  
Db 4453 GGACAGCAAGGCTGGGACAGGAAGTACATTTCTTGGAGGATCAAAAGTCTCTCATTTAT 4512  
Qy 4531 GACATGAAGCCAGAGAGCTTGAAGAGCCGTTGGAGAAATTTGAGCTGTGCTTCCC 4590  
Db 4513 GACATGAAGCCAGAGAGCTTGAAGAGCCGTTGGAGAAATTTGAGCTGTGCTTCCC 4572  
Qy 4591 GACGGGATGTATCTATTATCATGTCGCTTCCGAACTCGCAAAATACAGCCAAA 4650  
Db 4573 GACGGGATGTATCTATTATCATGTCGCTTCCGAACTCGCAAAATACAGCCAAA 4632  
Qy 4651 GCA----- 4653  
Db 4633 GAGATGTCCCATACATACTGAAGATGGAATCTCACCGCACACACCTGTGCGCCGG 4692  
Qy 4654 ----- 4656  
Db 4693 AGAACCTTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTACCGCCTTA 4752  
Qy 4654 ----- 4656  
Db 4753 GAATCAGTTGTCGAGGTGGGAGAGTTTCTAGGGAAGAGCAGAGCTGATGCTAAAATG 4812  
Qy 4681 CTTGGAATCCTCTGTAAGTGAAGTGTGACCTGTAGACATGAATGACGCTG 4740  
Db 4813 CTTGGAATCCTCTGTAAGTGAAGTGTGACCTGTAGACATGAATGACGCTG 4872  
Qy 4741 CCTTCAGTGACAGGTGTGTGTGGGACCGGAGAGGCTCTACGCCCTGAATGTC 4800  
Db 4873 CCTTCAGTGACAGGTGTGTGTGGGACCGGAGAGGCTCTACGCCCTGAATGTC 4932  
Qy 4801 TTGAAAACTCCTTAACCATGTCCAGGAATGGAGCGAGTCTTCAAAATTTATATTATC 4860  
Db 4933 TTGAAAACTCCTTAACCATGTCCAGGAATGGAGCGAGTCTTCAAAATTTATATTATC 4992  
Qy 4861 AAGGACCTGGAGAGTACTCATGTAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 4920  
Db 4993 AAGGACCTGGAGAGTACTCATGTAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAC 5052  
Qy 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCGCCAGCCGACATCTCA 4980  
Db 5053 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCGCCAGCCGACATCTCA 5112  
Qy 4981 CCCAACATTTTGAAGCTGTCAAGGCTGCACCTGTTGGGCGCAGGAAGATTGAGAAC 5040  
Db 5113 CCCAACATTTTGAAGCTGTCAAGGCTGCACCTGTTGGGCGCAGGAAGATTGAGAAC 5172  
Qy 5041 GGGCTCTGCATCTGTGACGCCATGCCAGCAAGTCTGTCATTCTCCGCTACAAACGAAAC 5100  
Db 5173 GGGCTCTGCATCTGTGACGCCATGCCAGCAAGTCTGTCATTCTCCGCTACAAACGAAAC 5232  
Qy 5101 CTCAGCAAAATCTGCATTCGGAAAGAGATAGACACCTCAGAGCCCTGACGCTGTATCCAC 5160  
Db 5233 CTCAGCAAAATCTGCATTCGGAAAGAGATAGACACCTCAGAGCCCTGACGCTGTATCCAC 5292  
Qy 5161 TTCACCAATTTACAGTATCTCTATTGAAACCAATTAATTTCTAGCAATCCACATGAAGCAG 5220  
Db 5293 TTCACCAATTTACAGTATCTCTATTGAAACCAATTAATTTCTAGCAATCCACATGAAGCAG 5352

RESULT 12  
ABQ78871

ID ABQ78871 standard; cdna; 5877 BP.

XX  
AC ABQ78871;

XX  
DT 10-OCT-2002 (first entry)

XX  
DE Human kinase cdna #2.

XX  
KW Human; kinase; enzyme; serine-threonine kinase; neotropic; cytostatic;  
KW Citron rho-interacting kinase; gene therapy; mental disorder; cancer;  
XX gene; ss.

OS Homo sapiens.

Qy 5221 TACACGCTCGAGGAATTCCTGGATAAAGATGACCATTCCTTGGACACTGCTGTGTTTGGC 5280  
Db 5353 TACACGCTCGAGGAATTCCTGGATAAAGATGACCATTCCTTGGACACTGCTGTGTTTGGC 5412  
Qy 5281 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGCGAGGAG 5340  
Db 5413 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGCGAGGAG 5472  
Qy 5341 GAGTACTGTGTGTTTCCACGAATTTGGAGTCTTCTGTGATTTTACGGAAGACCTAGC 5400  
Db 5473 GAGTACTGTGTGTTTCCACGAATTTGGAGTCTTCTGTGATTTTACGGAAGACCTAGC 5532  
Qy 5401 CGCACAGACATCTCAAGTGGAGTCCGTTACCTTTGGCCCTTTGCTTACAGAGAACCTAT 5460  
Db 5533 CGCACAGACATCTCAAGTGGAGTCCGTTACCTTTGGCCCTTTGCTTACAGAGAACCTAT 5592  
Qy 5461 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5520  
Db 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCCTCAGCA 5652  
Qy 5521 GGGACCCCTCCCGAGGCTACCTGGACATCCCGAAACCCGCGCTACCTGGGCCCTTGCATT 5580  
Db 5653 GGGACCCCTCCCGAGGCTACCTGGACATCCCGAAACCCGCGCTACCTGGGCCCTTGCATT 5712  
Qy 5581 TCCTCAGGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGGTCAATTTGCTGC 5640  
Db 5713 TCCTCAGGAGCGATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGGTCAATTTGCTGC 5772  
Qy 5641 AAGGGAAACCTCTGTGAAGAGTCCGGCACTGAAACACACCGGGGCCCTCCACTCCCGC 5700  
Db 5773 AAGGGAAACCTCTGTGAAGAGTCCGGCACTGAAACACACCGGGGCCCTCCACTCCCGC 5832  
Qy 5701 AGCAGCCCAACAGCGAGGCGCCACCCACGTAACAAGAGCACATCAACAGCGGTGGCC 5760  
Db 5833 AGCAGCCCAACAGCGAGGCGCCACCCACGTAACAAGAGCACATCAACAGCGGTGGCC 5892  
Qy 5761 TCACGCCAGCGCGCCGAGGCGCCAGCACCCGAGAGGCAAGCACACCCACCGC 5820  
Db 5893 TCACGCCAGCGCGCCGAGGCGCCAGCACCCGAGAGGCAAGCACACCCACCGC 5952  
Qy 5821 TACCGAGGGGCGGACCGAGCTGCGAGGACAAAGTCTCTGGCCGCCCTTGGAGCGA 5880  
Db 5953 TACCGAGGGGCGGACCGAGCTGCGAGGACAAAGTCTCTGGCCGCCCTTGGAGCGA 6012  
Qy 5881 GAGAAGTCCCGCCGCGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGCGCTGTT 5940  
Db 6013 GAGAAGTCCCGCCGCGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGCGCTGTT 6072  
Qy 5941 GAAGACAGCAGAGGGGCCGCTGCTGCGGAGCCCTGAGGACCCCGCTGTCCAGGTG 6000  
Db 6073 GAAGACAGCAGAGGGGCCGCTGCTGCGGAGCCCTGAGGACCCCGCTGTCCAGGTG 6132  
Qy 6001 AACAGGGAAGAGGCA 6017  
Db 6133 AACAGGGAAGGCA 6149

XX Key Location/Qualifiers  
FH 1..5877  
PT /\*tag= a  
FT /product= "Kinase"  
XX WO200259325-A2.  
XX 01-AUG-2002.  
XX 20-DEC-2001; 2001WO-US050497.  
XX 27-DEC-2000; 2000US-0258335P.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Yu X, Miranda M, Friddle CJ;  
XX WPI; 2002-599796/64.  
XX P-PSDB; ABB81928.  
XX Novel polynucleotide encoding human proteins that are structurally  
PT similar to animal kinases, useful for drug screening, diagnosis, in gene  
PT therapy of disorders and diseases e.g. cancer and pharmacogenomic  
PT applications.  
XX Disclosure; Page 44-45; 50pp; English.  
XX  
CC The invention relates to a novel human protein that shares structural  
CC similarity with animal kinases, including serine-threonine kinases,  
CC particularly Citron rho-interacting kinases. The proteins of the  
CC invention have nootropic and cytosolic activity. The polynucleotides may  
CC have a use in gene therapy. The encoded novel polypeptides are useful for  
CC generating antibodies, as reagents in diagnostic assays, for identifying  
CC other cellular gene products related to NHP and as reagents in assays for  
CC screening for compounds that are useful in the treatment of mental,  
CC biological or medical disorders and diseases including cancer. The  
CC sequence encodes a novel human kinase of the invention  
XX  
SQ Sequence 5877 BP; 1680 A; 1433 C; 1583 G; 1181 T; 0 U; 0 Other;  
Query Match 81.5%; Score 5358; DB 6; Length 5877;  
Best Local Similarity 95.8%; Pred. NO. 0;  
Matches 5634; Conservative 0; Mismatches 5; Indels 243; Gaps 3;  
QY 19 ATGTTGAAGTTCAGATGAGGCGCGGAATCTTTGGATGCTGGTCTGCTGAACCCATT 78  
DB 1 ATGTTGAAGTTCAGATGAGGCGCGGAATCTTTGGATGCTGGTCTGCTGAACCCATT 60  
QY 79 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA 138  
DB 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATGACTCAA 120  
QY 139 CAGCAGATGCTCTCTTTCCCGAGAAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198  
DB 121 CAGCAGATGCTCTCTTTCCCGAGAAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
QY 199 GAATGAGTCAGCTCTCTGATGAAGATTAAAGCAGTCGTCAGCAACTTTGTCGGGAAGTAT 258  
DB 181 GAATGAGTCAGCTCTCTGATGAAGATTAAAGCAGTCGTCAGCAACTTTGTCGGGAAGTAT 240  
QY 259 TCGACACCATAGCTGATTACAGGAGCTCCAGCCTTCGGCAAGAGACTTCGAAGTCAGA 318  
DB 241 TCGACACCATAGCTGATTACAGGAGCTCCAGCCTTCGGCAAGAGACTTCGAAGTCAGA 300  
QY 319 AGTCTTGTAGTGTGGTCACTTTCGTAAGTGCAGCTGGTGAAGAGAAACCCGGG 378  
DB 301 AGTCTTGTAGTGTGGTCACTTTCGTAAGTGCAGCTGGTGAAGAGAAACCCGGG 360  
QY 379 GACATCTATGCTATGAAGTGTAGAGAAAGAGGCTTTATTGGCCCGGAGCAGGTTTCA 438  
DB 361 GACATCTATGCTATGAAGTGTAGAGAAAGAGGCTTTATTGGCCCGGAGCAGGTTTCA 420

QY 439 TTTTGTGAGGAGGCGGAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 498  
DB TTTTGTGAGGAGGCGGAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 480  
QY 499 CAGTATGCCCTTTTCAGGACAAAAATCACCCTTTATCTGATGGAGGAATATCAGCTCGGAGGG 558  
DB CAGTATGCCCTTTTCAGGACAAAAATCACCCTTTATCTGATGGAGGAATATCAGCTCGGAGGG 540  
QY 559 GACTTGCTGTCACTTTTGAATAGATATGAGACCAAGTTAGATGAAAACTGATACAGTTT 618  
DB GACTTGCTGTCACTTTTGAATAGATATGAGACCAAGTTAGATGAAAACTGATACAGTTT 600  
QY 619 TACCTAGCTAGCTGATTTTGGCTGTTCAAGGCTTCACTGATGGGATACGTCGATCGA 678  
DB TACCTAGCTAGCTGATTTTGGCTGTTCAAGGCTTCACTGATGGGATACGTCGATCGA 660  
QY 679 GACATCAAGCTGAGAACATTTCTGTTGACCGGACAGGACACATCAAGCTGGTGATTT 738  
DB GACATCAAGCTGAGAACATTTCTGTTGACCGGACAGGACACATCAAGCTGGTGATTT 720  
QY 739 GGATCTCGCGGAAATGAATTTCAACCAAGATGTTGATCCCAAACTCCCGATTGGGACC 798  
DB GGATCTCGCGGAAATGAATTTCAACCAAGATGTTGATCCCAAACTCCCGATTGGGACC 780  
QY 799 CCAGATTACATGGCTCTGAAAGTGTGCTGATGACGCGGATGGAAGGACCTAC 858  
DB CCAGATTACATGGCTCTGAAAGTGTGCTGATGACGCGGATGGAAGGACCTAC 840  
QY 859 GGCCTGACCTGTGACTCGTGTGCTGAGTGGGCTGATTCGCTATGAGATGATTTATGGAGA 918  
DB GGCCTGACCTGTGACTCGTGTGCTGAGTGGGCTGATTCGCTATGAGATGATTTATGGAGA 900  
QY 919 TCCCGCTTCGAGAGGAACTCTGCCAGAACCTTCAATTAACATTTATGAATTTCCAGCGG 978  
DB TCCCGCTTCGAGAGGAACTCTGCCAGAACCTTCAATTAACATTTATGAATTTCCAGCGG 960  
QY 979 TTTTGTGAAATTTCCAGATGACCCCAAGTGACGAGTGTCTTCTGATCTGATCTCAAGC 1038  
DB TTTTGTGAAATTTCCAGATGACCCCAAGTGACGAGTGTCTTCTGATCTGATCTCAAGC 1020  
QY 1039 TTGTTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC 1098  
DB TTGTTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC 1080  
QY 1099 TCTAAATTCAGCTGGACACATTTGTAACCTCTCCCTCCCTCCCTCCCTCCCTCCCTCAAG 1158  
DB TCTAAATTCAGCTGGACACATTTGTAACCTCTCCCTCCCTCCCTCCCTCCCTCCCTCAAG 1140  
QY 1159 TCTGACGATGACACCTCCAAATTTTGAATGAACACAGAGAAGAAATTCGTGGGTTTCATCTCT 1218  
DB TCTGACGATGACACCTCCAAATTTTGAATGAACACAGAGAAGAAATTCGTGGGTTTCATCTCT 1200  
QY 1219 CCGTCCAGCTGAGCCCTCAGGCTCTCGGTTGAAGACTCCGTTTGTGGGGTTTTCG 1278  
DB CCGTCCAGCTGAGCCCTCAGGCTCTCGGTTGAAGACTCCGTTTGTGGGGTTTTCG 1260  
QY 1279 TACAGCAAGCACTGGGGATTTTGGTAGATCTGAGTCTGTTGTCGGGCTCTGGACTCC 1338  
DB TACAGCAAGCACTGGGGATTTTGGTAGATCTGAGTCTGTTGTCGGGCTCTGGACTCC 1320  
QY 1339 CTTGCCAAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1398  
DB CTTGCCAAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1380  
QY 1399 TCTCAGGACAAAGTGTCAACAGATGGAGCAAGAAATGACCCGGTTTACATCGGAGGTGTCA 1458  
DB TCTCAGGACAAAGTGTCAACAGATGGAGCAAGAAATGACCCGGTTTACATCGGAGGTGTCA 1440  
QY 1459 GAGTGGAGGCTGTCTTAGTCAGAGGAGGTGGAAGCTGAAGCCCTCTGAGACTCAGAGA 1518  
DB GAGTGGAGGCTGTCTTAGTCAGAGGAGGTGGAAGCTGAAGCCCTCTGAGACTCAGAGA 1500  
QY 1519 TCCCTCTCGGAGCAGGACCTTGCTACCTACATCACAGATGAGTAGCTTAAAGCGAAGT 1578

[illegible]

QY 3691 GAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGGCACATTCTCTCAA 3750  
|  
|  
|  
Db 3721 GAAACATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGGCACATTCTCTCAA 3780  
|  
|  
|  
QY 3751 CAACCAAACTCATTTCTGCAAGCCAAATGGACCAACTGCTCTAAAAAGAAAAAG 3810  
|  
|  
|  
Db 3781 CAACCAAACTCATTTCTGCAAGCCAAATGGACCAACTGCTCTAAAAAGAAA--- 3838  
|  
|  
|  
QY 3811 GGTTATTATGTCAGCGAAAGAGCCCTGCTTTACCCACACAGGTTCTCTGCGAGTAC 3870  
|  
|  
|  
Db 3839 -----AGTTCTCTGCGAGTAC 3855  
|  
|  
|  
QY 3871 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGACAGCTAGAGGAAGCC 3930  
|  
|  
|  
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGACAGCTAGAGGAAGCC 3915  
|  
|  
|  
QY 3931 CTTGAGAAACCCGATCGAGTCCGGTCCGCCCGGGAGGAAGCTGCCACCCGCAAGCA 3990  
|  
|  
|  
Db 3916 CTTGAGAAACCCGATCGAGTCCGGTCCGCCCGGGAGGAAGCTGCCACCCGCAAGCA 3975  
|  
|  
|  
QY 3991 ACGGACCAACCCACATCCACGCCAGCCACCGCAGGAGCAGATCGCCATGTCGCC 4050  
|  
|  
|  
Db 3976 ACGGACCAACCCACATCCACGCCAGCCACCGCAGGAGCAGATCGCCATGTCGCC 4035  
|  
|  
|  
QY 4051 ATCGTGGCGTGGCCAGAGCACCCAGCCAGTGCATGAGCCTGTGGCCCGCCATCCAGC 4110  
|  
|  
|  
Db 4036 ATCGTGGCGTGGCCAGAGCACCCAGCCAGTGCATGAGCCTGTGGCCCGCCATCCAGC 4095  
|  
|  
|  
QY 4111 CGCAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCCGGCTCTTAAGGAAAGCATGAC 4170  
|  
|  
|  
Db 4096 CGCAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCCGGCTCTTAAGGAAAGCATGAC 4155  
|  
|  
|  
QY 4171 CACAATATTCCTCACCGATTCACAGTGAACATCGAGGCCACAAAGTGTGCTGTG 4230  
|  
|  
|  
Db 4156 CACAATATTCCTCACCGATTCACAGTGAACATCGAGGCCACAAAGTGTGCTGTG 4215  
|  
|  
|  
QY 4231 TGTCTGATACCGTGACCTTTGAGCGCAGGACATCAAAATGCTCGAATGTCAAGTGATG 4290  
|  
|  
|  
Db 4216 TGTCTGATACCGTGACCTTTGAGCGCAGGACATCAAAATGCTCGAATGTCAAGTGATG 4275  
|  
|  
|  
QY 4291 TGTCAACCCAAAGTCTCCAGCTCTTCCAGCCACCTCGCGCTGTGCTGCTGAAATGTC 4350  
|  
|  
|  
Db 4276 TGTCAACCCAAAGTCTCCAGCTCTTCCAGCCACCTCGCGCTGTGCTGCTGAAATGTC 4335  
|  
|  
|  
QY 4351 ACACATTCACCGAGGCTTCTGCCGTGACAAATGAACTCCCGAGGTCTCCAGACCAAG 4410  
|  
|  
|  
Db 4336 ACACATTCACCGAGGCTTCTGCCGTGACAAATGAACTCCCGAGGTCTCCAGACCAAG 4395  
|  
|  
|  
QY 4411 GAGCCAGAGCAGCTTGACCTGGAAGGTTGGATGAGGTCGCCAGGATCAACAAAGCA 4470  
|  
|  
|  
Db 4396 GAGCCAGAGCAGCTTGACCTGGAAGGTTGGATGAGGTCGCCAGGATCAACAAAGCA 4455  
|  
|  
|  
QY 4471 GGAAGCAAGGCTGGACAGGAGTACATTGTCTGGAGGATCAAAAGTCTCTATTAT 4530  
|  
|  
|  
Db 4456 GGAAGCAAGGCTGGACAGGAGTACATTGTCTGGAGGATCAAAAGTCTCTATTAT 4515  
|  
|  
|  
QY 4531 GACAATGAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCCTGCC 4590  
|  
|  
|  
Db 4516 GACAATGAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCCTGCC 4575  
|  
|  
|  
QY 4591 GAGGGGATGTATCTATTATGCTGCGCTTGTGCTTCCGAATCGCAAAATACAGCCAAA 4650  
|  
|  
|  
Db 4576 GAGGGGATGTATCTATTATGCTGCGCTTGTGCTTCCGAATCGCAAAATACAGCCAAA 4635  
|  
|  
|  
QY 4651 GCA----- 4653  
|  
|  
|  
Db 4636 GCAGATGTCCCATACATCTGAAGATGGAATCTCACCCGACACACCTGTGCGCCGG 4695  
|  
|  
|  
QY 4654 ----- 4653  
|  
|  
|  
Db 4696 AGAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAGCGGTGAGGCTCACCGCTTA 4755

QY 4654 -----CAAAAGCAGAGCTGTGCTAAACTG 4680  
|  
|  
|  
Db 4756 GAATCAGTTGCGAGGTGGAGAGTTTCTAGGAAAGAGCAGAGCTGTGCTAAACTG 4815  
|  
|  
|  
QY 4681 CTTGAAACTCCCTGCTGAAACTGAAAGGTGATGACCGTCTAGACATGAACCTGCACGCTG 4740  
|  
|  
|  
Db 4816 CTTGAAACTCCCTGCTGAAACTGAAAGGTGATGACCGTCTAGACATGAACCTGCACGCTG 4875  
|  
|  
|  
QY 4741 CCCTTCAGTGCACAGGTGGTGTGGTGGCCACCGAGGAGGGCTCTACGCCCTGAATGTC 4800  
|  
|  
|  
Db 4876 CCCTTCAGTGCACAGGTGGTGTGGTGGCCACCGAGGAGGGCTCTACGCCCTGAATGTC 4935  
|  
|  
|  
QY 4801 TTGAAAACTCCCTTAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATC 4860  
|  
|  
|  
Db 4936 TTGAAAACTCCCTTAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATC 4995  
|  
|  
|  
QY 4861 AAGGACCTGGAGAGTACTCATGATAGCAGGAGAGAGCGGCACGTGTCTTGTGGAC 4920  
|  
|  
|  
Db 4996 AAGGACCTGGAGAGTACTCATGATAGCAGGAGAGAGCGGCACGTGTCTTGTGGAC 5055  
|  
|  
|  
QY 4921 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCCGACATCTCA 4980  
|  
|  
|  
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCCGACATCTCA 5115  
|  
|  
|  
QY 4981 CCCAACATTTTGAAGCTGCAAGGGTGCACCTTGTTCGGGAGGCAAGATTGGAAC 5040  
|  
|  
|  
Db 5116 CCCAACATTTTGAAGCTGCAAGGGTGCACCTTGTTCGGGAGGCAAGATTGGAAC 5175  
|  
|  
|  
QY 5041 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTCTCATTTCTCGCTACAAAGAAAC 5100  
|  
|  
|  
Db 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTCTCATTTCTCGCTACAAAGAAAC 5235  
|  
|  
|  
QY 5101 CTCAGCAAAATCTGCATCCGGAAGAGATAGACACCTCAGAGCCCTGCAGCTGTATCCAC 5160  
|  
|  
|  
Db 5236 CTCAGCAAAATCTGCATCCGGAAGAGATAGACACCTCAGAGCCCTGCAGCTGTATCCAC 5295  
|  
|  
|  
QY 5161 TTCAACCAATTTACAGTATCTCTANTGGAAACCAATAAATTTACGAATTCGACATGAAGCAG 5220  
|  
|  
|  
Db 5296 TTCACCAATTTACAGTATCTCTANTGGAAACCAATAAATTTACGAATTCGACATGAAGCAG 5355  
|  
|  
|  
QY 5221 TACACGCTCGAGGAATTCCTGGATGAATCAATTCCTTGGCACCTGTGTGTGTGCC 5280  
|  
|  
|  
Db 5356 TACACGCTCGAGGAATTCCTGGATGAATGAACATTCCTTGGCACCTGTGTGTGTGCC 5415  
|  
|  
|  
QY 5281 GCCTTCTTCCAAAGCTTCCCTGTCTCAATCGTGAAGTGAACAGCGGAGCGGAGAG 5340  
|  
|  
|  
Db 5416 GCCTTCTTCCAAAGCTTCCCTGTCTCAATCGTGAAGTGAACAGCGGAGCGGAGAG 5475  
|  
|  
|  
QY 5341 GAGTACTTGTGTGTGTTCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5400  
|  
|  
|  
Db 5476 GAGTACTTGTGTGTTCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5535  
|  
|  
|  
QY 5401 CGCAGACGATCTCAAGTGGAGTGTGCTTACCTTTGCTTTCCTTACAGAGAACCTCAT 5460  
|  
|  
|  
Db 5536 CGCAGACGATCTCAAGTGGAGTGTGCTTACCTTTGGCTTTCCTTACAGAGAACCTCAT 5595  
|  
|  
|  
QY 5461 CTGTTTCTGACCCACATTCACCTCAGTGAATTAATGAGATCCAGGACGCTCCTCAGCA 5520  
|  
|  
|  
Db 5596 CTGTTTCTGACCCACATTCACCTCAGTGAATTAATGAGATCCAGGACGCTCCTCAGCA 5555  
|  
|  
|  
QY 5521 GGAACCCCTCCCGAGCGTACCTGGACATCCGAAACCGCGCTTACCTGGGCGCTGCCATT 5580  
|  
|  
|  
Db 5556 GGAACCCCTCCCGAGCGTACCTGGACATCCGAAACCGCGCTTACCTGGGCGCTGCCATT 5715  
|  
|  
|  
QY 5581 TCCTCAGAGCGGATTTACTTTGGCGTCTCATACAGGATAAATTAAGGTTCATTCTGTC 5640  
|  
|  
|  
Db 5716 TCCTCAGAGCGGATTTACTTTGGCGTCTCATACAGGATAAATTAAGGTTCATTCTGTC 5775  
|  
|  
|  
QY 5641 AAGGAAACCTCTGTGAAGGAGTCCGGCAGTGAACACACCGGGGCGCTCCACCTCCCGC 5700  
|  
|  
|  
Db 5776 AAGGAAACCTCTGTGAAGGAGTCCGGCAGTGAACACACCGGGGCGCTCCACCTCCCGC 5835  
|  
|  
|  
QY 5701 AG 5702

CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
SQ Sequence 6609 BP; 1694 A; 1779 C; 1646 G; 1490 T; 0 U; 0 Other;  
  
Query Match 52.9%; Score 3475.2; DB 3; Length 6609;  
Best Local Similarity 94.2%; Pred. No. 0;  
Matches 3710; Conservative 0; Mismatches 33; Indels 195; Gaps 2;  
  
QY 2273 AAGTGTGGACAATCAGATAAAGAAAGACCTGCTGACAAGGAGACACTCGAACAATGA 2332  
Db |||||  
QY 2299 AGGTGTTGGACAATCAGATAAAGAAAGACCTGCTGACAAGGAGACACTCGAACAATGA 358  
Db |||||  
QY 2333 TGCAGACACGAGGAGGAGGCCCATGAGAAGGGCAAAATTCCTCAGCGAACAGAGCGCA 2392  
Db |||||  
QY 359 TGCAGACACGAGGAGGAGGCCCATGAGAAGGGCAAAATTCCTCAGCGAACAGAGCGCA 418  
Db |||||  
QY 2393 TGATCAATGTATGGATTCCCAAGTCCCTGGAACAGAGATTTGTGGAACTGTCGTG 2452  
Db |||||  
QY 419 TGATCAATGTATGGATTCCCAAGTCCCTGGAACAGAGATTTGTGGAACTGTCGTG 478  
Db |||||  
QY 2453 AAGCCATAAATTCGACCAATAGCAGTCTTTTACCCAAAGGAACATGAAGGCCCAAG 2512  
Db |||||  
QY 479 AAGCCATAAATTCGACCAATAGCAGTCTTTTACCCAAAGGAACATGAAGGCCCAAG 538  
Db |||||  
QY 2513 AAGGATGATTTCTGAATCAGGCAACAGAAATTTTACCTGGAGACACAGCTGGGAAGT 2572  
Db |||||  
QY 539 AAGGATGATTTCTGAATCAGGCAACAGAAATTTTACCTGGAGACACAGCTGGGAAGT 598  
Db |||||  
QY 2573 TGGAGGCCAGAACCCGAAAACCTGGAGGAGCAGCTGGAGAAGATCAGCCACCAAGACCACA 2632  
Db |||||  
QY 599 TGGAGGCCAGAACCCGAAAACCTGGAGGAGCAGCTGGAGAAGATCAGCCACCAAGACCACA 658  
Db |||||  
QY 2633 GTGCAAGAATTCGGCTCTGGAATCTGAGACACAGATTCGGGAGGTGAGTCTAGAGCAG 2692  
Db |||||  
QY 659 GTGCAAGAATTCGGCTCTGGAATCTGAGACACAGATTCGGGAGGTGAGTCTAGAGCAG 718  
Db |||||  
QY 2693 AGGAGCAGAACTGGAGCTCAAGGCCAGCTCAGAGACTACAGCTCTCCCTGCAGGAGC 2752  
Db |||||  
QY 719 AGGAGCAGAACTGGAGCTCAAGGCCAGCTCAGAGACTACAGCTCTCCCTGCAGGAGC 778  
Db |||||  
QY 2753 GCGAGTCACAGTTGACAGCCCTGCAGGCTGCAGGGCGGCCCTGGAGAGCCAGCTTCGCC 2812  
Db |||||  
QY 779 GCGAGTCACAGTTGACAGCCCTGCAGGCTGCAGGGCGGCCCTGGAGAGCCAGCTTCGCC 838  
Db |||||  
QY 2813 AGGCGAAGACAGAGCTGGAAAGAGACACAGAGAGCTGAAGAGGAGATCCAGGCATCA 2872  
Db |||||  
QY 839 AGGCGAAGACAGAGCTGGAAAGAGACACAGAGAGCTGAAGAGGAGATCCAGGCATCA 898  
Db |||||  
QY 2873 CGGCACATAGAGATGAATCCAGCGCAATTTGATGCTCTTCGTAAACAGCTGTACTGTAA 2932  
Db |||||  
QY 899 CGGCACATAGAGATGAATCCAGCGCAATTTGATGCTCTTCGTAAACAGCTGTACTGTAA 958  
Db |||||  
QY 2933 TCACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAGGACAACGCTGAATCTCAACAAAC 2992  
Db |||||  
QY 959 TCACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAGGACAACGCTGAATCTCAACAAAC 1018  
Db |||||  
QY 2993 AAAACTTCTACTTGTCCAAACAATCCATGAGCTTCTGCGGCCCAACGAGAGATTGTAC 3052  
Db |||||  
QY 1019 AAAACTTCTACTTGTCCAAACAATCCATGAGCTTCTGCGGCCCAACGAGAGATTGTAC 1078  
Db |||||  
QY 3053 AAATGCGAAGTGAAGTGAACCATCTCCGCGGGAGATCACCGAACCAGAGATGACGTTA 3112  
Db |||||  
QY 1079 AAATGCGAAGTGAAGTGAACCATCTCCGCGGGAGATCACCGAACCAGAGATGACGTTA 1138  
Db |||||  
QY 3113 CCAGCCAGAGCAACGATGGAGCTTGAAGACCAAGCTGACCATCTCGAGGAGAACAGG 3172  
Db |||||  
QY 1139 CCAGCCAGAGCAACGATGGAGCTTGAAGACCAAGCTGACCATCTCGAGGAGAACAGG 1198  
Db |||||  
QY 3173 TCATGGATTTGGAGGCCCTTAAACGATGAGTGTAGAAAAGAGCGGCACTGGAGGCCCT 3232  
Db |||||  
QY 1199 TCATGGATTTGGAGGCCCTTAAACGATGAGTGTAGAAAAGAGCGGCACTGGAGGCCCT 1258  
Db |||||

Db 5836 AG 5837  
  
RESULT 13  
AAC77568  
ID AAC77568 standard; cDNA; 6609 BP.  
XX  
AC AAC77568;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human ORFX ORF3123 polynucleotide sequence SEQ ID NO:6245.  
XX  
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
KW vulnary; antiparisonian; nootropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200058473-A2.  
XX  
XX 05-OCT-2000.  
XX  
XX 31-MAR-2000; 2000WO-US008621.  
XX  
XX 31-MAR-1999; 99US-0127607P.  
XX  
XX 02-APR-1999; 99US-0127636P.  
XX  
XX 05-APR-1999; 99US-0127728P.  
XX  
XX 30-MAR-2000; 2000US-00540763.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shinketsu RA, Leach M;  
XX  
XX WPI; 2000-602362/57.  
XX  
XX P-PSDB; AAB43359.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease.  
XX  
XX Claim 5; Page 5429-5433; 5507pp; English.  
XX  
XX AAC74446 to AAC77506 encode the proteins given in AAB40237 to AAB43397,  
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX  
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;  
XX antiparisonian; nootropic; neuroprotective; immunostimulant;  
XX anticonvulsant; antiarthritic; immunosuppressant; osteopathic;  
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
XX dermatological; immunosuppressive; antirheumatic; antithyroid; antianaemic. The  
XX antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The  
XX sequences can be used for determining the presence of or predisposition  
XX to, or preventing or treating pathological conditions associated with an  
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX  
XX proteins in gene therapy vectors. The proteins and nucleic acids may be  
XX used to treat cancers, proliferative disorders, neurodegenerative  
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
XX storage, systemic lupus erythematosus, severe combined immunodeficiency  
XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and





QY 1242 CTTCTCGGTGAAGAACTGCGGTTTGGGGTTTTGTTACAGCAAGCACTCGGGATTCT 1301  
Db 2416 CTTCTCGGTGAAGAACTGCGGTTTGGGGTTTTGTTACAGCAAGCACTCGGGATTCT 2357  
QY 1302 TGGTAGATCTGA ----- 1313  
Db 2356 TGGTAGATCTGAGACTTTTGGGGATGGATGAATACGTAAACACTTGGTCTAGTACC 2297  
QY 1314 -----GTCTGTGTGTCGGGTCTGACTCCCTGTCACAGACTAGC 1353  
Db 2296 TGGCACATGGAAAGCCCTTGGTGTGTGTCGGGTCTGACTCCCTGTCACAGACTAGC 2237  
QY 1354 TCCATGGAAAGAACTTCTCATCAAAAGCAAGAGAGCTACAAGACTCTCAGGACAAAGTGT 1413  
Db 2236 TCCATGGAAAGAACTTCTCATCAAAAGCAAGAGAGCTACAAGACTCTCAGGACAAAGTGT 2177  
QY 1414 CACAAGATGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTCAAGGTGGAGGTGTG 1473  
Db 2176 CACAAGATGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTCAAGGTGGAGGTGTG 2117  
QY 1474 CTTAGTCAGAGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGATCCCTCTCGGAGCAG 1533  
Db 2116 CTTAGTCAGAGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGATCCCTCTCGGAGCAG 2057  
QY 1534 GACCTTGCTACTACATCACAGAAATGACAGTGTAAAGCGAAGTTTGGAGCAAGCAGG 1593  
Db 2056 GACCTTGCTACTACATCACAGAAATGACAGTGTAAAGCGAAGTTTGGAGCAAGCAGG 1997  
QY 1594 ATGAGGTGTCAGAGAGGTGAAGCACTGACGCTTCCATCATATCAGAGAGCAG 1653  
Db 1996 ATGAGGTGTCAGAGAGGTGAAGCACTGACGCTTCCATCATATCAGAGAGCAG 1937  
QY 1654 AGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAGGCTCAAGTGAAGAAATGAGG 1713  
Db 1936 AGCCGGAAGCTCCAAGAAATCAAAGAGCAGGAGTACCAGGCTCAAGTGAAGAAATGAGG 1877  
QY 1714 TTGATGATGAATCAGTTGGAAGGATCTTGCTCAGCAAGAGAGCGAGTATCTCTAC 1773  
Db 1876 TTGATGATGAATCAGTTGGAAGGATCTTGCTCAGCAAGAGAGCGAGTATCTCTAC 1817  
QY 1774 GAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTAAGAAATCAAAGCGAAAGCGACAGAA 1833  
Db 1816 GAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTAAGAAATCAAAGCGAAAGCGACAGAA 1757  
QY 1834 TGTGAGATAACTGTTGAAGGCTAAGGATCAAGGGAAGCTGAAGTGGAGAAATGCG 1893  
Db 1756 TGTGAGATAACTGTTGAAGGCTAAGGATCAAGGGAAGCTGAAGTGGGAGAAATGCG 1697  
QY 1894 AAATGGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAGCTCCAAGAGAACTG 1953  
Db 1696 AAATGGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAGCTCCAAGAGAACTG 1637  
QY 1954 GAGAAGCT-----GCA 1965  
Db 1636 GAGAAGCTTAAAGCCAGCACGGAGGCCACCGAGCTGCTGCAGAAATATCCGCCAGGCA 1577  
QY 1966 AAGGAGCGCCAGAGGAGCTGGAGAGCTGCAGAACCGAGAGATCTTCTGAAGC 2025  
Db 1576 AAGGAGCGCCAGAGGAGCTGGAGAGCTGCAGAACCGAGAGATCTTCTGAAGC 1517  
QY 2026 ATCAGAAAGAGCTGGTGAAGCTGAGGAACCGCGCATTTCTTGAGAAACAAGTAAAG 2085  
Db 1516 ATCAGAAAGAGCTGGTGAAGCTGAGGAACCGCGCATTTCTTGAGAAACAAGTAAAG 1457  
QY 2086 AGACTAGAGCCATGAGCGGTAGAGAAACAGACTGAAGATGACATCCAGACAAATCC 2145  
Db 1456 AGACTAGAGCCATGAGCGGTAGAGAAACAGACTGAAGATGACATCCAGACAAATCC 1397  
QY 2146 CAAACAGATCAGCAGATGGCTGATAAAATTTCTGGAGCTGGAAGAGAAACATCGGAGGCC 2205  
Db 1396 CAAACAGATCAGCAGATGGCTGATAAAATTTCTGGAGCTGGAAGAGAAACATCGGAGGCC 1337  
QY 2206 CAAGTCTGCCCCAGCACCTAGAGTGCACCTGAAACAGAAAGAGCAGCAGTATGA-GGA 2264

Db 1336 CAAGTCTCAGCCAGCACCTAGAGTGCACTGAAACAGAAAGAGCAGCACTATGAGGGA 1277  
QY 2265 AAAGATTAAAGTTTGGCAATCAGATAAAGAAAGACTGTGCTGACAAAGGAGACACTGGA 2324  
Db 1276 AAAGATTAAAGTTTGGCAATCAGATAAAGAAAGACTGTGCTGACAAAGGAGACACTGGA 1217  
QY 2325 GAACATGATGAGAGACACAGAGGAGGAGCCCATGAGAGGGGCAAAATTTCTCAGCAACA 2384  
Db 1216 GAACATGATGAGAGACACAGAGGAGGAGCCCATGAGAGGGGCAAAATTTCTCAGCAACA 1157  
QY 2385 GAAGCGCATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAACAGAGGATTTGGA 2444  
Db 1156 GAAGCGCATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAACAGAGGATTTGGA 1097  
QY 2445 ACTGTCTGAAGCCCAATAAATTGACAGCAATATAGCAGTCTTTTACCCAAAGGAACATGAA 2504  
Db 1096 ACTGTCTGAAGCCCAATAAATTGACAGCAATATAGCAGTCTTTTACCCAAAGGAACATGAA 1037  
QY 2505 GGCCCAAGAGAGATGATTTCTGAACCTCAGCAACAGAAATTTTACCTGGAGACACAGGC 2564  
Db 1036 GGCCCAAGAGAGATGATTTCTGAACCTCAGCAACAGAAATTTTACCTGGAGACACAGGC 977  
QY 2565 TGGGAAGTTGGAGCCCGAGAACCCGAAACTGGAGGAGCAGCTGGAGAAAGATTCAGCCACCA 2624  
Db 976 TGGGAAGTTGGAGCCCGAGAACCCGAAACTGGAGGAGCAGCTGGAGAAAGATTCAGCCACCA 917  
QY 2625 AGACCAAGTCAAGAAATCGGCTGTGGAACCTGGAGACAGATTGGGGAGGTCACTCT 2684  
Db 916 AGACCAAGTCAAGAAATCGGCTGTGGAACCTGGAGACAGATTGGGGAGGTCACTCT 857  
QY 2685 AGACCAAGAGAGCAGAAACTGGAGCTCAAGCGCCAGCTCACAGAGTACAGCTCTCCCT 2744  
Db 856 AGACCAAGAGAGCAGAAACTGGAGCTCAAGCGCCAGCTCACAGAGTACAGCTCTCCCT 797  
QY 2745 GCAGGAGCGGAGTCAGATTGACAGCCCTGCAAGCTGCAAGGCGCCCTCGAGAGGCA 2804  
Db 796 GCAGGAGCGGAGTCAGATTGACAGCCCTGCAAGCTGCAAGGCGCCCTCGAGAGGCA 737  
QY 2805 GCTTCGCGAGCGAAGACAGAGCTGGAAGAGACACAGCAGAAAGCTGAAGAGGAGATCCA 2864  
Db 736 GCTTCGCGAGCGAAGACAGAGCTGGAAGAGACACAGCAGAAAGCTGAAGAGGAGATCCA 677  
QY 2865 GGCACTCAGGCAATAGAGATGAATCAGCGCAATTTGATGCTCTTCTGTAACAGCTG 2924  
Db 676 GGCACTCAGGCAATAGAGATGAATCAGCGCAATTTGATGCTCTTCTGTAACAGCTG 617  
QY 2925 TACTGTAATCAGACCTCGAGGAGCAGCTAAACCCAGCTCACCGAGGACAAACGCTGAAC 2984  
Db 616 TACTGTAATCAGACCTCGAGGAGCAGCTAAACCCAGCTCACCGAGGACAAACGCTGAAC 557  
QY 2985 CAAACAACAACTTCTACTTGTCCAAACACTCGATGAGGCTTTGCGCCCAACGAGCA 3044  
Db 556 CAAACAACAACTTCTACTTGTCCAAACACTCGATGAGGCTTTGCGCCCAACGAGCA 497  
QY 3045 GATTGTAACTCGAAGTGAAGTGAAGTCACTCCGCGGGAGATCAACGAAACGAGAT 3104  
Db 496 GATTGTAACTCGAAGTGAAGTGAAGTGAAGTCACTCCGCGGGAGATCAACGAAACGAGAT 437  
QY 3105 GCAGCTTACCAGCCAGCAAGCAACGATGGAGGCTCTGAAGACCAAGCTGCAACCTGCTGA 3164  
Db 436 GCAGCTTACCAGCCAGCAAGCAACGATGGAGGCTCTGAAGACCAAGCTGCAACCTGCTGA 377  
QY 3165 GGAACAGTCTGATGATGAGGCTTAAACGATGAGTCTGTAAGAAAGAGCGGAGTG 3224  
Db 376 GGAACAGTCTGATGATGAGGCTTAAACGATGAGTCTGTAAGAAAGAGCGGAGTG 317  
QY 3225 GGAGGCTGAGGAGGCTCTGCTGATGAGAAATCCAGTCTGAGTGTGCGGTTCGAGA 3284  
Db 316 GGAGGCTGAGGAGGCTCTGCTGATGAGAAATCCAGTCTGAGTGTGCGGTTCGAGA 257  
QY 3285 GCTGACAGATGCTGACACCCAGGAAACAGAGCAGGCGGAGGCGATCAGCGGATCAC 3344



QY 3415 CTCAAGAGCAGAGCTGAAGGCGGAGCCTCTCTGACAAAGCTCAATGACCTGGAGAAG 3474  
DB 541 CTCAAGAGCAGAGCTGAAGGCGGAGCCTCTCTGACAAAGCTCAATGACCTGGAGAAG 600  
QY 3475 AAGCATCTATGCTTGAATGAATGCCGAAGCTTACAGCAGAGCTGAGACTGAACGA 3534  
DB 601 AAGCATCTATGCTTGAATGAATGCCGAAGCTTACAGCAGAGCTGAGACTGAACGA 660  
QY 3535 GAGCTCAACACAGAGGCTTCTGGAAGAGCAGCAAGCAAAATTACAGCAGCAGATGGAACCTGCAG 3594  
DB 661 GAGCTCAACACAGAGGCTTCTGGAAGAGCAGCAAGCAAAATTACAGCAGCAGATGGAACCTGCAG 720  
QY 3595 AAAAATCACATTTTCGGTCTGAATCAAGGACTGCAAGAGCTTAGATCGGGCTGATCTA 3654  
DB 721 AAAAATCACATTTTCGGTCTGAATCAAGGACTGCAAGAGCTTAGATCGGGCTGATCTA 780  
QY 3655 CTGAAGACAGAAAGAGTGTGAGTATCAGCTGGAAGACATTCAGGTTCTCTATTCT 3714  
DB 781 CTGAAGACAGAAAGAGTGTGAGTATCAGCTGGAAGACATTCAGGTTCTCTATTCT 840  
QY 3715 CATGAAAGGTGAAATGGAAGCAGCTATTCTCAACAAACCAAACTCAATGATTCTG 3774  
DB 841 CATGAAAGGTGAAATGGAAGCAGCTATTCTCAACAAACCAAACTCAATGATTCTG 900  
QY 3775 CAAGCCAAATGACCAACCTGCTAAAGAAAGAGGGTTATTATTAGTCGACGGAAGAG 3834  
DB 901 CAAGCCAAATGACCAACCTGCTAAAGAAAGAGGGTTATTATTAGTCGACGGAAGAG 934  
QY 3835 GAOCCTGCTTTACCCACACAGAGTTCCTCTGCAATCAATGAGTGAAGCTGCGCCCTGGAG 3894  
DB 935 -----AGGTTCTCTGCAATCAATGAGTGAAGCTGCGCCCTGGAG 975  
QY 3895 AAGGAGAAAGCTCGTGTGAGAGCTAGAGAAAGCCCTTCAGAAAGCCGCAATCGAGTTC 3954  
DB 976 AAGGAGAAAGCTCGTGTGAGAGCTAGAGAAAGCCCTTCAGAAAGCCGCAATCGAGTTC 1035  
QY 3955 CGGTCCGCGGAGGAGAGTGCACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4014  
DB 1036 CGGTCCGCGGAGGAGAGTGCACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1095  
QY 4015 CCAGCCACCGGAGGAGCAGATGCGCATGTCGCCATCGTGGCTGCGGAGCAGACCCAG 4074  
DB 1096 CCAGCCACCGGAGGAGCAGATGCGCATGTCGCCATCGTGGCTGCGGAGCAGACCCAG 1155  
QY 4075 CCCAGTGCCATGAGCTGTCGCCCGCCCATCCAGCCGAGAAAGAGTCTCAACTCCA 4134  
DB 1156 CCCAGTGCCATGAGCTGTCGCCCGCCCATCCAGCCGAGAAAGAGTCTCAACTCCA 1215  
QY 4135 GAGGAATTTAGTCGGGCTTTAAGGAACGATGCACCAAAATTTCTCCACGATTTCAAC 4194  
DB 1216 GAGGAATTTAGTCGGGCTTTAAGGAACGATGCACCAAAATTTCTCCACGATTTCAAC 1275  
QY 4195 GTAGGACTGAATGCGGAGCAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA 4254  
DB 1276 GTAGGACTGAATGCGGAGCAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA 1335  
QY 4255 CGCAGGCAATCCAAATGTCTCGAATGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4314  
DB 1336 CGCAGGCAATCCAAATGTCTCGAATGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1395  
QY 4315 TTGCGAGCACTGCGGCTTGGCTGTGTAATGTGCAACAATTCACCGAGGCTTTGCG 4374  
DB 1396 TTGCGAGCACTGCGGCTTGGCTGTGTAATGTGCAACAATTCACCGAGGCTTTGCG 1455  
QY 4375 CGTGACAAATGAATCCCGAGTCTCCAGAGCAAGAGGCGCCAGCAGAGCTTGCACCTG 4434  
DB 1456 CGTGACAAATGAATCTCCCGAGTCTCCAGAGCAAGAGGCGCCAGCAGAGCTTGCACCTG 1515  
QY 4435 GAAGGTGATGAGTGTGCGGAGGAAATACAAACGAGGAGCAAGAGGCTTGGACAGGAAG 4494  
DB 1516 GAAGGTGATGAGTGTGCGGAGGAAATACAAACGAGGAGCAAGAGGCTTGGACAGGAAG 1575

QY 4495 TACATTGCTCTGGAGGATCAAAAGTCTCTCATTTATGACAAATGAAGCAGAGAGCTGGA 4554  
DB 1576 TACATTGCTCTGGAGGATCAAAAGTCTCTCATTTATGACAAATGAAGCAGAGAGCTGGA 1635  
QY 4555 CAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCCGACGGGAGTGTATCTATTCTATGCT 4614  
DB 1636 CAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCCGACGGGAGTGTATCTATTCTATGCT 1695  
QY 4615 GCGGTGTGCTTCCGAACTCGCAAAATACAGCCAAAGCA----- 4653  
DB 1696 GCGGTGTGCTTCCGAACTCGCAAAATACAGCCAAAGCAGATGTCCCATACATACTGAAG 1755  
QY 4654 ----- 4653  
DB 1756 ATGGAATCTCACCGGCACACACCTGCTGCGCGGAGAAACCTCTACTTTGCTAGTCCC 1815  
QY 4654 ----- 4653  
DB 1816 AGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTAGAATCAGTTGTCGAGGTGGAGA 1875  
QY 4654 -----GAAAAAGCAGAAAGTGTGTAATGCTTGGAAACCTCCTGCTGTAACCTG 4704  
DB 1876 GTTCTAGGGAAAAAGCAGAAAGTGTGTAATGCTTGGAAACCTCCTGCTGTAACCTG 1935  
QY 4705 GAAGGTGATGACCTGTAGACATGAACTGCACGCTGCCCTTCACTGACGAGGTGGTTG 4764  
DB 1936 GAAGGTGATGACCTGTAGACATGAACTGCACGCTGCCCTTCACTGACGAGGTGGTTG 1995  
QY 4765 GTGCGCCACCGAGGAAGGCTCTACGCCCTGAATGTCTTGAATACTCCCTAAACCTATGC 4824  
DB 1996 GTGCGCCACCGAGGAAGGCTCTACGCCCTGAATGTCTTGAATACTCCCTAAACCTATGC 2055  
QY 4825 CCAGAAATGAGAGAGTCTTCCAAATTTATATTATCAAGACCTGAGAGAGCTACTCATG 4884  
DB 2056 CCAGAAATGAGAGAGTCTTCCAAATTTATATTATCAAGACCTGAGAGAGCTACTCATG 2115  
QY 4885 ATACAGAGAGAGCGGCACTGTGTCTGTGGAAGTGAAGAAAGTGAAGTCCCTG 4944  
DB 2116 ATACAGAGAGAGCGGCACTGTGTCTGTGGAAGTGAAGAAAGTGAAGTCCCTG 2175  
QY 4945 GCCAGTCCCACTGCTGCCCGAGCCGCAATCTCAACCAAAATTTTGAAGCTGTCAAG 5004  
DB 2176 GCCAGTCCCACTGCTGCCCGAGCCGCAATCTCAACCAAAATTTTGAAGCTGTCAAG 2235  
QY 5005 GGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAAGGGCTCTGCACTCTGCGAGCATG 5064  
DB 2236 GGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAAGGGCTCTGCACTCTGCGAGCATG 2295  
QY 5065 CCCAGCAAGTGTGATTTCTCCGCTACAGCAAAACCTCAGCAAAATCTGCACTCCGAAA 5124  
DB 2296 CCCAGCAAGTGTGATTTCTCCGCTACAGCAAAACCTCAGCAAAATCTGCACTCCGAAA 2355  
QY 5125 GAGATGAGACCTCAGAGCCCTGCGCTGTATCCACTTCAACCAATTTACAGTATCTCAT 5184  
DB 2356 GAGATGAGACCTCAGAGCCCTGCGCTGTATCCACTTCAACCAATTTACAGTATCTCAT 2415  
QY 5185 GGAACCAATTAATTTCTACGAAATCGACATGAAGCAGTACACGCTCGAGGAATTTCTGGAT 5244  
DB 2416 GGAACCAATTAATTTCTACGAAATCGACATGAAGCAGTACACGCTCGAGGAATTTCTGGAT 2475  
QY 5245 AAGATGACCAATTTCTTGGCACTGCTGTGTTGGCGCTCTTCCACAGCTTCCCTGTC 5304  
DB 2476 AAGATGACCAATTTCTTGGCACTGCTGTGTTGGCGCTCTTCCACAGCTTCCCTGTC 2535  
QY 5305 TCAATCGTGAGGTGAACACGCGAGCGGAGAGGAGTACTTGTGTGTTTCCACGAA 5364  
DB 2536 TCAATCGTGAGGTGAACACGCGAGCGGAGAGGAGTACTTGTGTGTTTCCACGAA 2595  
QY 5365 TTTGGAGTGTGTTGAGTTCTTACGGAAGAGCTGAGCGGACAGACATCTCAAGTGGAGT 5424  
DB 2596 TTTGGAGTGTGTTGAGTTCTTACGGAAGAGCTGAGCGGACAGACATCTCAAGTGGAGT 2655  
QY 5425 CGCTTACCTTTGGCTTTGCTTACAGAAACCTATCTGTTGTGACCCACTTCAACTCA 5484

Db	2656	CGCTTACCTTTGGCCTTTGGCCTTACAGAGAACCCCTATCTGTTTGTGACCCACTTCAACTCA	2715
QY	5485	CTCGAAGTAATTGAGATCCAGGCACGCTCCTCAGCAGGACCCCTGCCGAGCGTACCTG	5544
Db	2716	CTCGAAGTAATTGAGATCCAGGCACGCTCCTCAGCAGGACCCCTGCCGAGCGTACCTG	2775
QY	5545	GACATCCCGAACCCCGCTACCTGGGCCCTGCCATTTCCTCAGGAGCGATTACTTGGCG	5604
Db	2776	GACATCCCGAACCCCGCTACCTGGGCCCTGCCATTTCCTCAGGAGCGATTACTTGGCG	2835
QY	5605	TCCTCATACCAGGATAAATTAGGGTCAATTTCCTCAGGAGGAACTCGTGAAGGAGTCC	5664
Db	2836	TCCTCATACCAGGATAAATTAGGGTCAATTTCCTCAGGAGGAACTCGTGAAGGAGTCC	2895
QY	5665	GGCACTGAACACACCGGGGCCCTCCACCTCCCGCAG	5702
Db	2896	GGCACTGAACACACCGGGGCCCTCCACCTCCCGCAG	2933

Search completed: July 3, 2004, 20:57:55  
Job time : 2262 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 11:30:09 ; Search time 15353 Seconds  
(without alignments)  
17387.465 Million cell updates/sec

Title: US-10-017-216-3  
Perfect score: 6159  
Sequence: 1 atgttgagtcacaaatagg.....ttctgagacagattattgc 6159

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues 6940544  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl :

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rat.\*
- 36: em.htg.nam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	6159	100.0	6159	6	AX429514 Sequence
2	6159	100.0	6574	6	AX429512 Sequence
3	5668	92.0	8603	6	AX671108 Sequence
4	5666.6	92.0	6156	6	AX671112 Sequence
5	5666.4	92.0	6298	6	AX504254 Sequence
6	5666.2	92.0	6165	6	AX671105 Sequence
7	5661.4	91.9	6165	6	AX574425 Sequence
8	5650.4	91.7	6159	6	AX166510 Sequence
9	5631.8	91.4	6189	6	AX503780 Sequence
10	5629.8	91.4	6201	6	AX503778 Sequence
11	5534.8	89.9	8576	9	AX257469 Homo sapi
12	5358	87.0	5877	6	AX574427 Sequence
13	4627	75.1	6954	10	AF086824 Mus muscu
14	3602.6	58.5	5952	10	AF039218 Rattus no
15	3593.8	58.4	4967	10	AF070066 Mus muscu
16	3546.4	57.6	5019	10	MMU39904 U39904 Mus musculu
17	2436.8	39.6	5261	6	AX671110 Sequence
18	2436.8	39.6	5261	9	AB023166 Homo sapi
19	2187.4	35.5	3495	10	BC051185 BC051185 Mus muscu
20	2187.4	35.5	3506	10	BC023775 BC023775 Mus muscu
21	1395	22.6	1485	6	AX671037 Sequence
22	1395	22.6	1765	6	AX671044 Sequence
23	1393.4	22.6	2066	6	AX642956 Sequence
24	1391.8	22.6	1515	6	AR253937 Sequence
25	1082.4	17.6	2380	6	AX671042 Sequence
26	1082.4	17.6	2380	10	AF086823 AF086823 Mus muscu
27	1074.4	17.4	1386	10	AF070085 AF070085 Rattus no
28	750	12.2	1133	6	AX411948 Sequence
29	648	10.5	1058	6	AX053315 Sequence
30	592	9.6	817	6	AX053416 Sequence
31	434	7.0	2284	9	AK123136 Homo sapi
32	350.4	5.7	354	6	BD059649 BD059649 Secreted
33	311.4	5.1	136436	9	AC002563 AC002563 Human PAC
34	311.4	5.1	147862	9	AC079317 AC079317 Homo sapi
35	311.4	5.1	190514	2	AC026363 AC026363 Homo sapi
36	311.4	5.1	208686	9	AC079262 AC079262 Homo sapi
37	258	4.2	258	6	AR269950 Sequence
38	257	4.2	32658	2	AC135099 AC135099 Rattus no
39	257	4.2	234312	2	AC133614 AC133614 Rattus no
40	257	4.2	248557	2	AC097432 AC097432 Rattus no
41	256.2	4.2	207404	2	AC115795 AC115795 Mus muscu
42	250	4.1	254	6	BD071469 BD071469 Secreted
43	244	4.0	3835	6	AX817259 Sequence
44	244	4.0	3985	6	AX817261 Sequence
45	243.6	4.0	1530	6	AX504235 Sequence

ALIGNMENTS

RESULT 1  
AX429514  
LOCUS AX429514 6159 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 3 from Patent WO0234896.  
ACCESSION AX429514  
VERSION AX429514.1 GI:21540792  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Kapeller-Libermann,R.  
TITLE 13245, a novel human myotonic dystrophy type protein kinase and  
uses therefor

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0234896-A 3 02-MAY-2002;  
MILLENNIUM PHARM INC (US)

FEATURES  
source  
Location/Qualifiers  
1..6159  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 6159; DB 6; Length 6159;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTTCAAAATATGGAGCGCGGAATCCTTTTGGATGCTGGTCTGCTGAACCCATT 60  
DB 1 ATGTTGAAGTTCAAAATATGGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAACCCATT 60

QY 61 GCAGCGGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCCCTTTATGACTCAA 120  
DB 61 GCAGCGGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCCCTTTATGACTCAA 120

QY 121 CAGCAGATGCTCCTCTTCCCGAGAGGGAATAGATGCCCTCTTGTCTCTTTGAA 180  
DB 121 CAGCAGATGCTCCTCTTCCCGAGAGGGAATAGATGCCCTCTTGTCTCTTTGAA 180

QY 181 GAATGCAATCAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAAGTAT 240  
DB 181 GAATGCAATCAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAAGTAT 240

QY 241 TCCGACACCATAGCTCAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAGTCAGA 300  
DB 241 TCCGACACCATAGCTCAGTTACAGGAGCTCCAGCTTCGCGAAAGGACTTCGAAGTCAGA 300

QY 301 AGTCTTGTAGTTGGTGCATCTTGTGAGTGCAGTGGTGAAGAGAAAGCAACCGGG 360  
DB 301 AGTCTTGTAGTTGGTGCATCTTGTGAGTGCAGTGGTGAAGAGAAAGCAACCGGG 360

QY 361 GACATCTGCTATGAAGTGAAGAGAGGCTTTATTTGGCCAGGAGAGCTTCA 420  
DB 361 GACATCTGCTATGAAGTGAAGAGAGGCTTTATTTGGCCAGGAGAGCTTCA 420

QY 421 TTTTGTGAGGAAGCGGCAATATATCTCGAAGCACAGCCGCTGGATCCCCAAATTA 480  
DB 421 TTTTGTGAGGAAGCGGCAATATATCTCGAAGCACAGCCGCTGGATCCCCAAATTA 480

QY 481 CAGTATGCCCTTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGCTCGAGGG 540  
DB 481 CAGTATGCCCTTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGCTCGAGGG 540

QY 541 GACTTGTGTCATTTTGAATAGATATGAGGACAGTTAGATGAACCTGATACAGTTT 600  
DB 541 GACTTGTGTCATTTTGAATAGATATGAGGACAGTTAGATGAACCTGATACAGTTT 600

QY 601 TACTAGCTAGCTGATTTTGGCTTTCACAGGTTTCATCTGATGGGATACCTGATCGA 660  
DB 601 TACTAGCTAGCTGATTTTGGCTTTCACAGGTTTCATCTGATGGGATACCTGATCGA 660

QY 661 GACATCAAGCTTCAGAACATCTCTGTTGACCGACAGGACACATCAAGCTGTGGATT 720  
DB 661 GACATCAAGCTTCAGAACATCTCTGTTGACCGACAGGACACATCAAGCTGTGGATT 720

QY 721 GGAATCTGCCGCGAAAATGAATCAAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780  
DB 721 GGAATCTGCCGCGAAAATGAATCAAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780

QY 781 CCAGATTACATGCTCTGAGTCTGACTGTGATGAACCGGGATCGAAAGCCACTAC 840  
DB 781 CCAGATTACATGCTCTGAGTCTGACTGTGATGAACCGGGATCGAAAGCCACTAC 840

QY 841 GGCTGGACTGTGACTGGTGGTCAAGTGGGCGTGAATTCGCTATGAGATGATTTATGGGAGA 900  
DB 841 GGCTGGACTGTGACTGGTGGTCAAGTGGGCGTGAATTCGCTATGAGATGATTTATGGGAGA 900

QY 901 TCCCCCTTCGAGAGGGAACCTCTGCCAGAACCTTCAATAACATTAATGAATTTCCAGCGG 960  
DB 901 TCCCCCTTCGAGAGGGAACCTCTGCCAGAACCTTCAATAACATTAATGAATTTCCAGCGG 960

QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATCAAGC 1020  
DB 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATCAAGC 1020

QY 1021 TTGTTGCGGCCAGAAAAGAGAGACTGAAGTTTGAAGTCTTCTGCTGCCATCTCTTCTTC 1080  
DB 1021 TTGTTGCGGCCAGAAAAGAGAGACTGAAGTTTGAAGTCTTCTGCTGCCATCTCTTCTTC 1080

QY 1081 TCTAAAATTGACTGGAAACAATTCGTAATCTCTCTCCCTTCGTTCCACCTCAAG 1140  
DB 1081 TCTAAAATTGACTGGAAACAATTCGTAATCTCTCTCCCTTCGTTCCACCTCAAG 1140

QY 1141 TCTGACGATGACCTCCAAATTTTGTGAACCAAGAGAAATTCGTGGGTTTCACTCTCT 1200  
DB 1141 TCTGACGATGACCTCCAAATTTTGTGAACCAAGAGAAATTCGTGGGTTTCACTCTCT 1200

QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1260  
DB 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1260

QY 1261 TACAGCAAGCACTGGGGATCTTGGTAGATCTGAGTCTGTTGTCGGGTCGGACTCC 1320  
DB 1261 TACAGCAAGCACTGGGGATCTTGGTAGATCTGAGTCTGTTGTTGTCGGGTCGGACTCC 1320

QY 1321 CCTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATAAAAAGCAAGAGCTACAAGAC 1380  
DB 1321 CCTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATAAAAAGCAAGAGCTACAAGAC 1380

QY 1381 TCTCAGCAAGTGTCAAAAGTGAAGAGCAAAATGACCCGTTTACATCCGAGAGTCA 1440  
DB 1381 TCTCAGCAAGTGTCAAAAGTGAAGAGCAAAATGACCCGTTTACATCCGAGAGTCA 1440

QY 1441 GAGTGTGAGGCTGTCTTAGTCAGAGGAGTGGAGTGAAGGCTCTGAGACTCAGAGA 1500  
DB 1441 GAGTGTGAGGCTGTCTTAGTCAGAGGAGTGGAGTGAAGGCTCTGAGACTCAGAGA 1500

QY 1501 TCCCTCTCGAGCAGGACCTTGTCTA CTTA CAT CACAGATGCAAGTGAAGTGAAGT 1560  
DB 1501 TCCCTCTCGAGCAGGACCTTGTCTA CTTA CAT CACAGATGCAAGTGAAGTGAAGT 1560

QY 1561 TTGAGCAAGCAGGATGGAGGTGCCAGAGAGTGCACAAACACCTGCAAGCTTCTCCAT 1620  
DB 1561 TTGAGCAAGCAGGATGGAGGTGCCAGAGAGTGCACAAACACCTGCAAGCTTCTCCAT 1620

QY 1621 GATATCAGAGCAGAGCCCGGAAGCTTCAAAGAAATCAAAGAGCAGGAGTACCAGGCTCAA 1680  
DB 1621 GATATCAGAGCAGAGCCCGGAAGCTTCAAAGAAATCAAAGAGCAGGAGTACCAGGCTCAA 1680

QY 1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
DB 1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740

QY 1741 CGGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800  
DB 1741 CGGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800

QY 1801 CGGAAACCGACAGATGTACGATTAACCTTTGAAGCTTAAGGATCAAGGAGGCTGAA 1860  
DB 1801 CGGAAACCGACAGATGTACGATTAACCTTTGAAGCTTAAGGATCAAGGAGGCTGAA 1860

QY 1861 GTGGGAGAAATATGGAATCTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
DB 1861 GTGGGAGAAATATGGAATCTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920

QY 1921 CTCCAAAGAGAAATCTGGAGAGGCTGCAAGAGCAGGAGGAGGAGCTGGAGAGCTG 1980  
DB 1921 CTCCAAAGAGAAATCTGGAGAGGCTGCAAGAGCAGGAGGAGGAGCTGGAGAGCTG 1980

QY 1981 CAGAACCGAGGAGTCTTCTCTGAGGAGCATCAGAAAAGAGCTGCTGGAAGCTGAGGAACGC 2040

1981 CAGAACCGAGAGGATTCCTCTGAGGCATCAGAAAGAGCTGGTGAAGCTCAGGAACGC 2040  
2041 CCGCATTCCTCGAGAAACAGGTAAAGAGACTAGAGACCAATGGAGCGTAGAGAAACAGA 2100  
2041 CCGCATTCCTCGAGAAACAGGTAAAGAGACTAGAGACCAATGGAGCGTAGAGAAACAGA 2100  
2101 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGT 2160  
2101 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGT 2160  
2161 GAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTCAGCCCCAGCACCCTAGAGAGTGACCTG 2220  
2161 GAGCTCGAAGAGAAACATCGGAGGCCCAAGTCTCAGCCCCAGCACCCTAGAGAGTGACCTG 2220  
2221 AAACAGAAAGACGACGACCTATCAGGAAAGATTAAGTGTGGAACAATCAGATAAAGAA 2280  
2221 AAACAGAAAGACGACGACCTATCAGGAAAGATTAAGTGTGGAACAATCAGATAAAGAA 2280  
2281 GACCTGGCTGACAAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGAGGCCCAT 2340  
2281 GACCTGGCTGACAAGGAGACACTGGAGAACATGATGCAGAGACACGAGGAGAGGCCCAT 2340  
2341 GAGAGGCGCAAAATTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCAGATC 2400  
2341 GAGAGGCGCAAAATTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCAGATC 2400  
2401 AGATCCCTGGAAACAGAGATTCGTGAACTGTCTGAAGCCAAATAAATTCGACGAAATAGC 2460  
2401 AGATCCCTGGAAACAGAGATTCGTGAACTGTCTGAAGCCAAATAAATTCGACGAAATAGC 2460  
2461 AGTCTTTTACCCAAAGAAACATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520  
2461 AGTCTTTTACCCAAAGAAACATGAAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520  
2521 CAGAAATTTTACTCGAGACACAGCTGGAGAGTTGGAGGCCCCAGAACCGAAACCTGGAG 2580  
2521 CAGAAATTTTACTCGAGACACAGCTGGAGAGTTGGAGGCCCCAGAACCGAAACCTGGAG 2580  
2581 GAGCAGCTGGAGAAGATCAGCCACCAAGACCACTGACAGAAATCGGCTCGTGAACCTG 2640  
2581 GAGCAGCTGGAGAAGATCAGCCACCAAGACCACTGACAGAAATCGGCTCGTGAACCTG 2640  
2641 GAGACAAGATTCGGGGAGTCACTTAGAGACGAGGAGCAGAACTGGAGCTCAGAGCC 2700  
2641 GAGACAAGATTCGGGGAGTCACTTAGAGACGAGGAGCAGAACTGGAGCTCAGAGCC 2700  
2701 CAGCTCACAGAGCTACAGCTCTCCCTGAGGAGCGGAGTCACTGACAGCCCTGCGAG 2760  
2701 CAGCTCACAGAGCTACAGCTCTCCCTGAGGAGCGGAGTCACTGACAGCCCTGCGAG 2760  
2761 GCTGACGGGGCGGCTGAGAGCCAGCTTCGCCAGGCGAGACAGAGCTGGAAAGACC 2820  
2761 GCTGACGGGGCGGCTGAGAGCCAGCTTCGCCAGGCGAGACAGAGCTGGAAAGACC 2820  
2821 ACAGCAGAAGCTGAAGAGAGATCCAGGACCTCAGGCACATAGAGATGAATCCAGGCG 2880  
2821 ACAGCAGAAGCTGAAGAGAGATCCAGGACCTCAGGCACATAGAGATGAATCCAGGCG 2880  
2881 AAATTGTGCTCTTCGTAAACAGCTGTACTGTAATCAAGAGCTGGAGGAGCAGCTAAAC 2940  
2881 AAATTGTGCTCTTCGTAAACAGCTGTACTGTAATCAAGAGCTGGAGGAGCAGCTAAAC 2940  
2941 CAGCTGACCGAGCAACGCTGAACCTCAACACCAAACTTCTACTTGTCCAAACAACTC 3000  
2941 CAGCTGACCGAGCAACGCTGAACCTCAACACCAAACTTCTACTTGTCCAAACAACTC 3000  
3001 GATGAGGCTTCTGGCGCCCAACAGCAGATTTGACAACTGCGAAGTGAAGTGGACCATCTC 3060  
3001 GATGAGGCTTCTGGCGCCCAACAGCAGATTTGACAACTGCGAAGTGAAGTGGACCATCTC 3060  
3061 CCGCGGAGATCAGCGAACAGAGATGCAAGTTTACGCCAGAGCAACCAAGTAGGAGGCT 3120

3061 CGCCGGAGATCAGGAAACGAGAGATGACAGCTTACAGCCAGAAGCAACGATGGAGGCT 3120  
3121 CTGAAGACCACTGTCACCATGCTGGAGGAACAGGTCTATGATTTGGAGGCCCTTAAACGAT 3180  
3121 CTGAAGACCACTGTCACCATGCTGGAGGAACAGGTCTATGATTTGGAGGCCCTTAAACGAT 3180  
3181 GAGCTGCTAGAAAAAGAGCGGAGCTGGAGGCGCTGGAGGAGCGCTCTGGGTGATGAGAAA 3240  
3181 GAGCTGCTAGAAAAAGAGCGGAGCTGGAGGCGCTGGAGGAGCGCTCTGGGTGATGAGAAA 3240  
3241 TCCAGATTTGAGTGTCCGGTTCCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGC 3300  
3241 TCCAGATTTGAGTGTCCGGTTCCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGC 3300  
3301 AGGCGAGAGCCGATCAGCGGATCACCAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG 3360  
3301 AGGCGAGAGCCGATCAGCGGATCACCAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG 3360  
3361 GAGCACAAGGCTGAGATTTCCGTCTCGACAGCGCTCTCAAGAGCAGAAAGCTGAAGGCC 3420  
3361 GAGCACAAGGCTGAGATTTCCGTCTCGACAGCGCTCTCAAGAGCAGAAAGCTGAAGGCC 3420  
3421 GAGAGCCTCTCTGACAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAAT 3480  
3421 GAGAGCCTCTCTGACAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAAT 3480  
3481 GCCCGAAGCTTACAGCAGAAAGCTGGAGACTGAACGAGAGCTCAACACAGAGGCTTCTGGAA 3540  
3481 GCCCGAAGCTTACAGCAGAAAGCTGGAGACTGAACGAGAGCTCAACACAGAGGCTTCTGGAA 3540  
3541 GAGCAAGCCAAATACAGCAGAGATGGAGCTCGAGNAAAATCACATTTTCCCTCTGACT 3600  
3541 GAGCAAGCCAAATACAGCAGAGATGGAGCTCGAGNAAAATCACATTTTCCCTCTGACT 3600  
3601 CAAGCAGTCAAGAAAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3660  
3601 CAAGCAGTCAAGAAAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3660  
3661 GAGTATCAGCTGGAAAACATTCAGGTTCTCTATCTCATGAAAAGTGAAGTGAAGGC 3720  
3661 GAGTATCAGCTGGAAAACATTCAGGTTCTCTATCTCATGAAAAGTGAAGTGAAGGC 3720  
3721 ACTATTTCTCAACAAACCAAACTCATTGATTTCTCGACGCCAAATGGACCAACCTGCT 3780  
3721 ACTATTTCTCAACAAACCAAACTCATTGATTTCTCGACGCCAAATGGACCAACCTGCT 3780  
3781 AAAAAAGAAAAGGTTTATTTAGTCGACGGAAAGAGGACCCCTCTTTACCCACACAGGTT 3840  
3781 AAAAAAGAAAAGGTTTATTTAGTCGACGGAAAGAGGACCCCTCTTTACCCACACAGGTT 3840  
3841 CCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTCAGAG 3900  
3841 CCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAAAGAGAAAGCTCGCTGTCAGAG 3900  
3901 CTAGAGAAAGCCCTTCAGAAAGACCCGATCGAGCTCCGGTCCCGCCGGAGGAAGTGGCC 3960  
3901 CTAGAGAAAGCCCTTCAGAAAGACCCGATCGAGCTCCGGTCCCGCCGGAGGAAGTGGCC 3960  
3961 CACCGAAAAGCAACGGACCAACCCATCCAGCCAGCCACCGCGGAGGAGCAGATC 4020  
3961 CACCGAAAAGCAACGGACCAACCCATCCAGCCAGCCACCGCGGAGGAGCAGATC 4020  
4021 GCCATGTCGCCCATCGTGGGTGCGCAGAGCACCCAGTCCCATGAGCCTGCTGGCC 4080  
4021 GCCATGTCGCCCATCGTGGGTGCGCAGAGCACCCAGTCCCATGAGCCTGCTGGCC 4080  
4081 CCGCCATCCAGCGCGAGAAAGGAGTCTCAACTCCAGAGGAAATTTAGTCGGCGCTTTAAG 4140  
4081 CCGCCATCCAGCGCGAGAAAGGAGTCTCAACTCCAGAGGAAATTTAGTCGGCGCTTTAAG 4140  
4141 GAACGATGACACCAATATTTCTCACCAGTTCACGTAAGGACTGAACATGCGAGCCACA 4200  
4141 GAACGATGACACCAATATTTCTCACCAGTTCACGTAAGGACTGAACATGCGAGCCACA 4200

QY 4201 AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGACGCCAGGCATCCAAAATGTCTCGAA 4260  
DB 4201 AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGACGCCAGGCATCCAAAATGTCTCGAA 4260  
QY 4261 TGTCAAGTATGTGTCAACCCCAAGTGTCTCAAGTGTCTTCCAGGCAACCTGTGGGCTTGGCT 4320  
DB 4261 TGTCAAGTATGTGTCAACCCCAAGTGTCTCAAGTGTCTTCCAGGCAACCTGTGGGCTTGGCT 4320  
QY 4321 GCTGAATATGCCACACATTTACCGAGGCTTCTGCCGTGACAAATGAATCCCAAGT 4380  
DB 4321 GCTGAATATGCCACACATTTACCGAGGCTTCTGCCGTGACAAATGAATCCCAAGT 4380  
QY 4381 CTCAGACCAAGAGGCCCGCAGCAGCTTTGCCTGGAAAGGTGGATGAAGTGGCCAGG 4440  
DB 4381 CTCAGACCAAGAGGCCCGCAGCAGCTTTGCCTGGAAAGGTGGATGAAGTGGCCAGG 4440  
QY 4441 AATAACAAACGAGGACAGCAAGCTTGGACAGGAGTACATTTCTCTGGAGGATCAAAA 4500  
DB 4441 AATAACAAACGAGGACAGCAAGCTTGGACAGGAGTACATTTCTCTGGAGGATCAAAA 4500  
QY 4501 GTCTCTCATTTATGACAAATGAAGCCAGAGAGTGGACAGAGCCGGTGGAAATTTGAG 4560  
DB 4501 GTCTCTCATTTATGACAAATGAAGCCAGAGAGTGGACAGAGCCGGTGGAAATTTGAG 4560  
QY 4561 CTGTGCTTCCCGACCGGGATGTATCTATTCATGTGCGCTTGGTCTTCGAACTCGCA 4620  
DB 4561 CTGTGCTTCCCGACCGGGATGTATCTATTCATGTGCGCTTGGTCTTCGAACTCGCA 4620  
QY 4621 AATACAGCAAAACGAGAAAGCAGAGCTGTATCTAATCTGTTGGAACTCCCTGTG 4680  
DB 4621 AATACAGCAAAACGAGAAAGCAGAGCTGTATCTAATCTGTTGGAACTCCCTGTG 4680  
QY 4681 AAATCTGGAAGTGAATGACCGCTAGACATGAATGACATGACCGCTTCTAGTACCAAGTG 4740  
DB 4681 AAATCTGGAAGTGAATGACCGCTAGACATGAATGACATGACCGCTTCTAGTACCAAGTG 4740  
QY 4741 GTCTTGGTGGACCCAGAGAGGCTCTAGCGCTGAATGCTTGAATACTCCCTAACC 4800  
DB 4741 GTCTTGGTGGACCCAGAGAGGCTCTAGCGCTGAATGCTTGAATACTCCCTAACC 4800  
QY 4801 CATGTCCCAGGAATGGAGCAGTCTTCCAAATTTATATTAATCAAGACCTGGAGAAGTGA 4860  
DB 4801 CATGTCCCAGGAATGGAGCAGTCTTCCAAATTTATATTAATCAAGACCTGGAGAAGTGA 4860  
QY 4861 CTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAGTGAAGAACTGAAACAG 4920  
DB 4861 CTCATGATAGCAGGAGAGAGCGGCACTGTGTCTTGTGGAGTGAAGAACTGAAACAG 4920  
QY 4921 TCCTTGGCCAGTCCCACTGCTGCTGCCAGCCGACATCTCACCCAAATTTTGAAGCT 4980  
DB 4921 TCCTTGGCCAGTCCCACTGCTGCTGCCAGCCGACATCTCACCCAAATTTTGAAGCT 4980  
QY 4981 GTCAAGGCTGCCATTTGTTGGGAGGAGAGATTTGAGAACGGGCTGTGATCTGTGCA 5040  
DB 4981 GTCAAGGCTGCCATTTGTTGGGAGGAGAGATTTGAGAACGGGCTGTGATCTGTGCA 5040  
QY 5041 GCCATGCCAGCAAGTCTGCTATCTCCGCTACACGAAACCTCAGCAAAATCTGCAATC 5100  
DB 5041 GCCATGCCAGCAAGTCTGCTATCTCCGCTACACGAAACCTCAGCAAAATCTGCAATC 5100  
QY 5101 CGGAAAGAGATAGACCTCAGAGCCCTGAGCTGATCCATCTCACCAATTTACAGTATC 5160  
DB 5101 CGGAAAGAGATAGACCTCAGAGCCCTGAGCTGATCCATCTCACCAATTTACAGTATC 5160  
QY 5161 CTCATTTGGACCAATTAATTTCTACGAAATCGCATGAACGATGACGCTCGAGAAATTC 5220  
DB 5161 CTCATTTGGACCAATTAATTTCTACGAAATCGCATGAACGATGACGCTCGAGAAATTC 5220  
QY 5221 CTGGATAAGAAATGACCACTTCTTGGACCTGTGTGTTTGGCCGCTTCTTCCAAACAGCTTC 5280  
DB 5221 CTGGATAAGAAATGACCACTTCTTGGACCTGTGTGTTTGGCCGCTTCTTCCAAACAGCTTC 5280

QY 5281 CCTGTCTCAATCGTGTGAGGTGAACAGCGCAGGGCAGGAGAGTACTTGTGTGTTC 5340  
DB 5281 CCTGTCTCAATCGTGTGAGGTGAACAGCGCAGGGCAGGAGAGTACTTGTGTGTTC 5340  
QY 5341 CAGAAATTCGAGTGTTCGTGGATTTCTTACGGAAGAGTACGCGCACAGCATCTCAAG 5400  
DB 5341 CAGAAATTCGAGTGTTCGTGGATTTCTTACGGAAGAGTACGCGCACAGCATCTCAAG 5400  
QY 5401 TGGAGTGCCTTACCTTTGGCTTTGGCTACAGAGAACCTTATCTGTTTGTGACCCATTC 5460  
DB 5401 TGGAGTGCCTTACCTTTGGCTTTGGCTACAGAGAACCTTATCTGTTTGTGACCCATTC 5460  
QY 5461 AACTCATTCAAGTAAATTTAGATCCAGGACGCTCTCAGCAGGACCTTCCGCGAGG 5520  
DB 5461 AACTCATTCAAGTAAATTTAGATCCAGGACGCTCTCAGCAGGACCTTCCGCGAGG 5520  
QY 5521 TACTGTGACATCCCGAACCCGCTACCTGGGCTTCCCATTTCTCAGGAGGATTTAC 5580  
DB 5521 TACTGTGACATCCCGAACCCGCTACCTGGGCTTCCCATTTCTCAGGAGGATTTAC 5580  
QY 5581 TTGGCGTCTCATACAGGATAAATTTAGGCTCATTTGCTGCAAGGAAACCTCTGTAAG 5640  
DB 5581 TTGGCGTCTCATACAGGATAAATTTAGGCTCATTTGCTGCAAGGAAACCTCTGTAAG 5640  
QY 5641 GAGTCCGGCACTGAACACACCCGGGCGCTCCACCTCCCGCAGCAGCCCAACAGCGA 5700  
DB 5641 GAGTCCGGCACTGAACACACCCGGGCGCTCCACCTCCCGCAGCAGCCCAACAGCGA 5700  
QY 5701 GGCCCAACCCAGTAAACAGGACATCAACAGCGGTGGCTTCCAGCCAGCGCGCC 5760  
DB 5701 GGCCCAACCCAGTAAACAGGACATCAACAGCGGTGGCTTCCAGCCAGCGCGCC 5760  
QY 5761 GAGGCGCCAGCCAGCAGGACCAAGCAGCAGCCCGCTACCGGAGGCGGAC 5820  
DB 5761 GAGGCGCCAGCCAGCAGGACCAAGCAGCAGCCCGCTACCGGAGGCGGAC 5820  
QY 5821 GAGTGTGCGAGGACAGTCTCTGCGCGCCCTGGAGCAGAGAGTCCCGCGCGG 5880  
DB 5821 GAGTGTGCGAGGACAGTCTCTGCGCGCCCTGGAGCAGAGAGTCCCGCGCGG 5880  
QY 5881 ATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTTTGAAAGACAGCAGCGG 5940  
DB 5881 ATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTTTGAAAGACAGCAGCGG 5940  
QY 5941 CGGTGTGCTGCGGAGCGGTGAGGACCCCGCTGTCCAGGTGAACAGGAGGAGGCG 6000  
DB 5941 CGGTGTGCTGCGGAGCGGTGAGGACCCCGCTGTCCAGGTGAACAGGAGGAGGCG 6000  
QY 6001 AGTGCCTCTCAAGTTCACGGTTACACTGTCACTTATGATGGAATAAAGCTG 6060  
DB 6001 AGTGCCTCTCAAGTTCACGGTTACACTGTCACTTATGATGGAATAAAGCTG 6060  
QY 6061 GACAACTGCCAGTAACTGTGCTGAGTCACTCAGATCACTGAGTGAATGAGAAATCCGG 6120  
DB 6061 GACAACTGCCAGTAACTGTGCTGAGTCACTCAGATCACTGAGTGAATGAGAAATCCGG 6120  
QY 6121 CAGCAGTGTGAAAGTCTGTTCTGAGAACAGATTTATTC 6159  
DB 6121 CAGCAGTGTGAAAGTCTGTTCTGAGAACAGATTTATTC 6159

## RESULT 2

AX429512 LOCUS 6574 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 1 from Patent WO0234896.  
ACCESSION AX429512  
VERSION AX429512.1 GI:21540791  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1

AUTHORS Kapeller-Libermann,R.  
TITLE 13245, a novel human myotonic dystrophy type protein kinase and  
uses therefor  
JOURNAL Patent: WO 0234896-A 1 02-MAY-2002;  
MILLENNIUM PHARM INC (US)  
FEATURES Location/Qualifiers  
Source 1..5574  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 6159; DB 6; Length 6574;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGTGTTGGTGTGCTGCTCAACCCATT 60  
Db 1.9 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGTGTTGGTGTGCTGCTCAACCCATT 78  
QY 61 GCACGCGGCGCTCCAGCTGTAATCTGTTCCAGGGGAACCAACCTTTATGACTCAA 120  
Db 79 GCACGCGGCGCTCCAGCTGTAATCTGTTCCAGGGGAACCAACCTTTATGACTCAA 138  
QY 121 CAGCAGATGTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
Db 139 CAGCAGATGTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198  
QY 181 GAATCGAGTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCACTTTGTCCGGAAGTAT 240  
Db 199 GAATCGAGTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCACTTTGTCCGGAAGTAT 258  
QY 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGGCAAGGACTTCGAAGTCAGA 300  
Db 259 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCGGCAAGGACTTCGAAGTCAGA 318  
QY 301 AGTCTGTAGTTGGTCACTTTGCTGAAGTCAGGTGGTAAAGAGAGAAACCAACCGGG 360  
Db 319 AGTCTGTAGTTGGTCACTTTGCTGAAGTCAGGTGGTAAAGAGAGAAACCAACCGGG 378  
QY 361 GACATCTAGCTATGAAGTGTAGAGAGAGGCTTTATTGGCCCGAGGACGAGTTTCA 420  
Db 379 GACATCTAGCTATGAAGTGTAGAGAGAGGCTTTATTGGCCCGAGGACGAGTTTCA 438  
QY 421 TTTTGTGAGGAAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Db 439 TTTTGTGAGGAAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 498  
QY 481 CAGTATGCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCTGGAGGG 540  
Db 499 CAGTATGCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCTGGAGGG 558  
QY 541 GACTTGCTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAACCTGATACAGTTT 600  
Db 559 GACTTGCTCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAAACCTGATACAGTTT 618  
QY 601 TACCTAGCTGAGCTGATTTTGCTGTTTCACAGCTTCATCTGATGGGATACGTCATCGA 660  
Db 619 TACCTAGCTGAGCTGATTTTGCTGTTTCACAGCTTCATCTGATGGGATACGTCATCGA 678  
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCCGACAGGACACATCAAGCTGGTGAATTT 720  
Db 679 GACATCAAGCCTGAGAACATTTCTGTTGACCCGACAGGACACATCAAGCTGGTGAATTT 738  
QY 721 GGATCTGCCCGGAAATGAATTAACCAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780  
Db 739 GGATCTGCCCGGAAATGAATTAACCAAGATGGTGAATGCCAAACTCCCGATTGGGACC 798  
QY 781 CCAGATTACATGGCTCCTGAAGTGTGACTGTGATGAACGGGGATGGAAGGACCTTAC 840  
Db 799 CCAGATTACATGGCTCCTGAAGTGTGACTGTGATGAACGGGGATGGAAGGACCTTAC 858  
QY 841 GGCCTGGAGTGTGACTGGTGGTCACTGGGCGGTGATTGCCCTATGAGATGATTTATGGGAGA 900

Db 859 GGCCTGGAGTGTGACTGGTGGTCAGTGGGCGTGATTCCTTATGAGATGATTTATGGGAGA 918  
QY 901 TCCCCCTTCGACAGAGGGAACCTCTGCGCAGAACCTTCAATAAACATTATGAATTTCCACGG 960  
Db 919 TCCCCCTTCGACAGAGGGAACCTCTGCGCAGAACCTTCAATAAACATTATGAATTTCCACGG 978  
QY 961 TTTTGAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 979 TTTTGAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038  
QY 1021 TTGTTGTGGCGCCAGAAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCTCTTTCTTC 1080  
Db 1039 TTGTTGTGGCGCCAGAAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTGCCATCTCTTTCTTC 1098  
QY 1081 TCTAAATTTGACTGGAACAACATTGCTGTAACCTCTCTCCCCCTTGGTTCACACCTCAAG 1140  
Db 1099 TCTAAATTTGACTGGAACAACATTGCTGTAACCTCTCTCCCCCTTGGTTCACACCTCAAG 1158  
QY 1141 TCTGACCATCACACCTCCAAATTTTGTATGAACACAGAGAGAAATTCGTGGGTTTCATCTCT 1200  
Db 1159 TCTGACCATCACACCTCCAAATTTTGTATGAACACAGAGAGAAATTCGTGGGTTTCATCTCT 1218  
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCGTTTGGGGTTTTCG 1260  
Db 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCGTTTGGGGTTTTCG 1278  
QY 1261 TACAGCAAGCACTGGGATTTCTTGGTAGATCTGAGTCTGTTGTCGGCTCTGGAATCC 1320  
Db 1279 TACAGCAAGCACTGGGATTTCTTGGTAGATCTGAGTCTGTTGTCGGCTCTGGAATCC 1338  
QY 1321 CTTGCCAAGACTAGCTCCATCGAAGAAAGAACTTCTCATCAAAAGCAAAAGAGCTACAAGAC 1380  
Db 1339 CTTGCCAAGACTAGCTCCATCGAAGAAAGAACTTCTCATCAAAAGCAAAAGAGCTACAAGAC 1398  
QY 1381 TCTCAGACAAGTGTACAAAGATGGACAGGAATAGCCCCGTTTACATCGAGAGTGTCA 1440  
Db 1399 TCTCAGACAAGTGTACAAAGATGGACAGGAATAGCCCCGTTTACATCGAGAGTGTCA 1458  
QY 1441 GAGTGTGAGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
Db 1459 GAGTGTGAGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1518  
QY 1501 TCCTCTCTGAGCAGGACCTTGTCTACATCAAGAAATGCAAGTACGAGTAAAGCGAAGT 1560  
Db 1519 TCCTCTCTGAGCAGGACCTTGTCTACATCAAGAAATGCAAGTACGAGTAAAGCGAAGT 1578  
QY 1561 TTGGAGCAAGCAGGATGGAGGTGTCCTCAGAGAGGTGACAAAGCACTGCGAGCTTCTCAT 1620  
Db 1579 TTGGAGCAAGCAGGATGGAGGTGTCCTCAGAGAGGTGACAAAGCACTGCGAGCTTCTCAT 1638  
QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAAGCAGGAGTACCAGGCTCAA 1680  
Db 1639 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAAGCAGGAGTACCAGGCTCAA 1698  
QY 1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
Db 1699 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758  
QY 1741 CGGAGTCTATCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1800  
Db 1759 CGGAGTCTATCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1818  
QY 1801 CGGAAACGCAAGATGTCAGCATAACTTTGAAAGCTTAAGGATCAAGGAGAGCTGAA 1860  
Db 1819 CGGAAACGCAAGATGTCAGCATAACTTTGAAAGCTTAAGGATCAAGGAGAGCTGAA 1878  
QY 1861 GTGGGAATATGCAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
Db 1879 GTGGGAATATGCAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938  
QY 1921 CTCACAGAGAACTGGAGAGAGCTGCAAGAGGACGAGCGAGAGGAGCTGGAGAGAGCTG 1980

Db 1939 CTCRAAGAAACTGGAGRAGGCTGCAAAAGGCGAGCGAGGAGGAGCTGGAGAGCTG 1998  
Qy 1981 CAGAAACCGAGAGATTCTTCTGAGGATCATGAAAGAGCTGGTGAAGCTGAGAAAGC 2040  
Db 1999 CAGAAACCGAGAGATTCTTCTGAGGATCATGAAAGAGCTGGTGAAGCTGAGAAAGC 2058  
Qy 2041 CGCATTCTCTGAGAACAGGTAAAGAGACTAGACACCATGGAGCGTAGAGAAACAGA 2100  
Db 2059 CGCATTCTCTGAGAACAGGTAAAGAGACTAGACACCATGGAGCGTAGAGAAACAGA 2118  
Qy 2101 CTGAAGGATGACATCCAGACAAAATCCCAACAGATCCAGACATGGCTGATAAAATTCG 2160  
Db 2119 CTGAAGGATGACATCCAGACAAAATCCCAACAGATCCAGACATGGCTGATAAAATTCG 2178  
Qy 2161 GAGCTCGAAGAAACATCGGAGGCGCAAGTCTGAGCCAGCATCTAGAAAGTGCACCTG 2220  
Db 2179 GAGCTCGAAGAAACATCGGAGGCGCAAGTCTGAGCCAGCATCTAGAAAGTGCACCTG 2238  
Qy 2221 AAACAGAAAGAGCAGCAGCTATGAGGAAAGATTAAAGTGTGGACAATCAGATAAAGAAA 2280  
Db 2239 AAACAGAAAGAGCAGCAGCTATGAGGAAAGATTAAAGTGTGGACAATCAGATAAAGAAA 2298  
Qy 2281 GAGCTGGCTGACAGAGACACTGGAGAACATGATCGACAGACACAGAGAGAGGCCCAT 2340  
Db 2299 GAGCTGGCTGACAGAGACACTGGAGAACATGATCGACAGACACAGAGAGAGGCCCAT 2358  
Qy 2341 GAGAAAGGCAAAATCTCAGCGAAACAGAAAGGCGATGATCAATGCTATGGATTCCAAGATC 2400  
Db 2359 GAGAAAGGCAAAATCTCAGCGAACAGAAAGGCGATGATCAATGCTATGGATTCCAAGATC 2418  
Qy 2401 AGATCCCTGAAACAGAGGATTGTGGAACCTGTCTGAAGCCAAATAAATTCGACCAATAGC 2460  
Db 2419 AGATCCCTGAAACAGAGGATTGTGGAACCTGTCTGAAGCCAAATAAATTCGACCAATAGC 2478  
Qy 2461 AGTCTTTTACCAAGGAACATGAGGCGCCAGAGAGATGATTTCTGAACCTCAGGCAA 2520  
Db 2479 AGTCTTTTACCAAGGAACATGAGGCGCCAGAGAGATGATTTCTGAACCTCAGGCAA 2538  
Qy 2521 CAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCCCAGAACCGAAATCTGGAG 2580  
Db 2539 CAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCCCAGAACCGAAATCTGGAG 2598  
Qy 2581 GAGCAGCTGGAGAGATCAGCCACCAAGACCAAGTGAAGAAATCGGCTGTGGAATG 2640  
Db 2599 GAGCAGCTGGAGAGATCAGCCACCAAGACCAAGTGAAGAAATCGGCTGTGGAATG 2658  
Qy 2641 GAGACAAGATTGGGGAGGTCAAGTCTAGAGACAGAGGAGCAAAATCTGAGCTCAAGGCG 2700  
Db 2659 GAGACAAGATTGGGGAGGTCAAGTCTAGAGACAGAGGAGCAAAATCTGAGCTCAAGGCG 2718  
Qy 2701 CAGCTCACAGAGTACAGCTCTCCCTGAGAGCGCGAGTCAAGTTCAGACCCCTGAG 2760  
Db 2719 CAGCTCACAGAGTACAGCTCTCCCTGAGAGCGCGAGTCAAGTTCAGACCCCTGAG 2778  
Qy 2761 GCTGCAAGGCGGCGCTGAGAGCGAGTTCGCGAGCGAAGACAGAGCTGGAAGAGACC 2820  
Db 2779 GCTGCAAGGCGGCGCTGAGAGCGAGTTCGCGAGCGAAGACAGAGCTGGAAGAGACC 2838  
Qy 2821 ACAGCAGAAGCTGAAGAGAGATCCAGGCACTCAAGGCACTAGAGATGAATTCAGGCG 2880  
Db 2839 ACAGCAGAAGCTGAAGAGAGATCCAGGCACTCAAGGCACTAGAGATGAATTCAGGCG 2898  
Qy 2881 AAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCACAAGACTTGAGAGAGCTAAAC 2940  
Db 2899 AAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCACAAGACTTGAGAGAGCTAAAC 2958  
Qy 2941 CAGCTGACCGAGACAACTGAACTCAACCAAACTTCTACTTTGTCCTCAAACTC 3000  
Db 2959 CAGCTGACCGAGACAACTGAACTCAACCAAACTTCTACTTTGTCCTCAAACTC 3018  
Qy 3001 GATGAGGCTTCGGCCCAACGACGAGATTGTACAACTCGGAAGTGAAGTGGACCATCTC 3060  
Db 3019 GATGAGGCTTCGGCCCAACGACGAGATTGTACAACTCGGAAGTGAAGTGGACCATCTC 3078

Qy 3061 CGCCGGAGATCAGGAACGAGAGATCGAGTTTACCGAGCAGAGCAAAACGATGGAGCT 3120  
Db 3079 CGCCGGAGATCAGGAACGAGAGATCGAGTTTACCGAGCAGAGCAAAACGATGGAGCT 3138  
Qy 3121 CTGAAGACCAAGTGCACCATCTGGAGGAAACAGGTTCATGATTTGGAGGCCCTTAAACGAT 3180  
Db 3139 CTGAAGACCAAGTGCACCATCTGGAGGAAACAGGTTCATGATTTGGAGGCCCTTAAACGAT 3198  
Qy 3181 GAGTGTCTAGAAAAGAGCGGAGTGGAGGCGCTGGAGGCGCTCTGGTGTATGAGAAA 3240  
Db 3199 GAGTGTCTAGAAAAGAGCGGAGTGGAGGCGCTGGAGGCGCTCTGGTGTATGAGAAA 3258  
Qy 3241 TCCAGATTTGAGTGTGCGGTTTCGAGAGCTCGAGAGATGCTGACACCGAGAAACAGAGC 3300  
Db 3259 TCCAGATTTGAGTGTGCGGTTTCGAGAGCTCGAGAGATGCTGACACCGAGAAACAGAGC 3318  
Qy 3301 AGGCGAGAGCCGATCAGCCGATCAACGAGTCTGCGCAGGTGTGAGCTGGCAGTGAAG 3360  
Db 3319 AGGCGAGAGCCGATCAGCCGATCAACGAGTCTGCGCAGGTGTGAGCTGGCAGTGAAG 3378  
Qy 3361 GAGCACAGGCTGAGATTCTCGCTCTGCGCAGGCTCTCAAGAGCAGAGCTGAAGGCC 3420  
Db 3379 GAGCACAGGCTGAGATTCTCGCTCTGCGCAGGCTCTCAAGAGCAGAGCTGAAGGCC 3438  
Qy 3421 GAGAGCCTCTCTGACAGGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAT 3480  
Db 3439 GAGAGCCTCTCTGACAGGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAT 3498  
Qy 3481 GCCCGAGACTTACAGCAGAGCTGGAGACTGAGCAGAGCTCAACAGAGGCTTCTGGAA 3540  
Db 3499 GCCCGAGACTTACAGCAGAGCTGGAGACTGAGCAGAGCTCAACAGAGGCTTCTGGAA 3558  
Qy 3541 GAGCAAGCCAAATTTACAGCAGCAGATGGAACCTCGAGAAAAATCACATTTTCCGCTGACT 3600  
Db 3559 GAGCAAGCCAAATTTACAGCAGCAGATGGAACCTCGAGAAAAATCACATTTTCCGCTGACT 3618  
Qy 3601 CAAGGATCGCAAGAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGATGACTTG 3660  
Db 3619 CAAGGATCGCAAGAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGATGACTTG 3678  
Qy 3661 GAGTATCAGCTGGAACACATTCAGGTTCTATTTCTCATGAAGGTGAATTTGAAGGC 3720  
Db 3679 GAGTATCAGCTGGAACACATTCAGGTTCTATTTCTCATGAAGGTGAATTTGAAGGC 3738  
Qy 3721 ACTATTTCTCAACAAACCAAACTCATTTGTTTCTGCAAGCCAAAATGGAACCACTGCT 3780  
Db 3739 ACTATTTCTCAACAAACCAAACTCATTTGTTTCTGCAAGCCAAAATGGAACCACTGCT 3798  
Qy 3781 AAAAAGAAAAGGTTTATTTAGTCCGAGGAAAGAGACCTCTTTTACCCACACAGGTT 3840  
Db 3799 AAAAAGAAAAGGTTTATTTAGTCCGAGGAAAGAGACCTCTTTTACCCACACAGGTT 3858  
Qy 3841 CCTCTGAGTACAATGAGCTGGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAG 3900  
Db 3859 CCTCTGAGTACAATGAGCTGGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAG 3918  
Qy 3901 CTAGAGAAAGCCCTTCAGAAAGACCGCATCGAGTCCGCTCGGTCGCGGAGGAGGAGCTGC 3960  
Db 3919 CTAGAGAAAGCCCTTCAGAAAGACCGCATCGAGTCCGCTCGGTCGCGGAGGAGGAGCTGC 3978  
Qy 3961 CACCGCAAAAGCAACGAGACCAACCAACCTCCACGCGACCCCGGAGGAGCAGATC 4020  
Db 3979 CACCGCAAAAGCAACGAGACCAACCAACCTCCACGCGACCCCGGAGGAGCAGATC 4038  
Qy 4021 GCCATGTCCGCATCTGTGCGGTTCGACAGAGCACAGCCCATGCCATGAGCTGCTGGCC 4080  
Db 4039 GCCATGTCCGCATCTGTGCGGTTCGACAGAGCACAGCCCATGCCATGAGCTGCTGGCC 4098  
Qy 4081 CCGGCATCCAGCGCAGAAAGAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTCTTAAG 4140  
Db 4099 CCGGCATCCAGCGCAGAAAGAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTCTTAAG 4158



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Zhu, Z.  
TITLE Regulation of human citron rho/rac-interacting kinase  
JOURNAL Patent: WO 0304523-A 4 16-JAN-2003;  
Bayer Aktiengesellschaft (DE)  
Location/Qualifiers

## FEATURES

source  
1..8603  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 92.0%; Score 5668; DB 6; Length 8603;  
Best Local Similarity 95.5%; Pred. No. 0;  
Matches 5962; Conservative 0; Mismatches 35; Indels 243; Gaps 3;  
QY 1 ATGTTGAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGGTCTGCTGAACCCATT 60  
Db |||||  
QY 1 ATGTTGAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGGTCTGCTGAACCCATT 60  
Db |||||  
QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAACCCACCCTTTATGACTCAA 120  
Db |||||  
QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAACCCACCCTTTATGACTCAA 120  
Db |||||  
QY 121 CAGCAGATGTCCTCTTCCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
Db |||||  
QY 121 CAGCAGATGTCCTCTTCCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
Db |||||  
QY 181 GAATGAGTCAGGCTGCTCTGATGAAGATTAAAGCAGTCAGCAACTTTCTCCGGAAGTAT 240  
Db |||||  
QY 181 GAATGAGTCAGGCTGCTCTGATGAAGATTAAAGCAGTCAGCAACTTTCTCCGGAAGTAT 240  
Db |||||  
QY 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAAAGGACTTCGAAGTCAGA 300  
Db |||||  
QY 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAAAGGACTTCGAAGTCAGA 300  
Db |||||  
QY 301 AGCTTTGTAGTTGCTGCTCACTTTGCTGAAGTCGAGTGTGAAGAGAAACCAACCGG 360  
Db |||||  
QY 301 AGCTTTGTAGTTGCTGCTCACTTTGCTGAAGTCGAGTGTGAAGAGAAACCAACCGG 360  
Db |||||  
QY 361 GACATCTATGCTATGAAAGTGAATGAAGAGAGAGGCTTTATTTGGCCACGAGCAGGTTTCA 420  
Db |||||  
QY 361 GACATCTATGCTATGAAAGTGAATGAAGAGAGAGGCTTTATTTGGCCACGAGCAGGTTTCA 420  
Db |||||  
QY 421 TTTTTCGAGAGAGCGGACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Db |||||  
QY 421 TTTTTCGAGAGAGCGGACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Db |||||  
QY 481 CAGTATGCTCTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540  
Db |||||  
QY 481 CAGTATGCTCTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540  
Db |||||  
QY 541 GACTTCTGCTACTTTTGAATAGATAGAGGACGATTTAGATGAAGAACTTGATCAGTTT 600  
Db |||||  
QY 541 GACTTCTGCTACTTTTGAATAGATAGAGGACGATTTAGATGAAGAACTTGATCAGTTT 600  
Db |||||  
QY 601 TACCTAGCTGAGTGAATTTGGCTGTTTCACAGCTTCATCTGATGGGATACGTGATCGA 660  
Db |||||  
QY 601 TACCTAGCTGAGTGAATTTGGCTGTTTCACAGCTTCATCTGATGGGATACGTGATCGA 660  
Db |||||  
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCCGACAGGACATCAAGCTGGTGGATTTT 720  
Db |||||  
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCCGACAGGACATCAAGCTGGTGGATTTT 720  
Db |||||  
QY 721 GGATCTGCCCGGAAATGAATTCACACAGATGGTGAATGCCCAACTCCGATTTGGGACC 780  
Db |||||  
QY 721 GGATCTGCCCGGAAATGAATTCACACAGATGGTGAATGCCCAACTCCGATTTGGGACC 780  
Db |||||  
QY 781 CCAGATTACATGGCTCTTGAAGTGTGATGATGAACCGGGATGGAAGGACCTTAC 840  
Db |||||  
QY 781 CCAGATTACATGGCTCTTGAAGTGTGATGATGAACCGGGATGGAAGGACCTTAC 840  
Db |||||

QY 841 GGCCTGGACTGTGACTGTGTGCTCAGTGGGCTGATTGCTATGAGATGATTATTATGGAGA 900  
Db |||||  
QY 841 GGCCTGGACTGTGACTGTGTGCTCAGTGGGCTGATTGCTATGAGATGATTATTATGGAGA 900  
Db |||||  
QY 901 TCCCTTCGCGAGAGGAACTCTGCGAGAACCTTCAATTAACATTAATTAATTAATTAATTA 960  
Db |||||  
QY 901 TCCCTTCGCGAGAGGAACTCTGCGAGAACCTTCAATTAACATTAATTAATTAATTAATTA 960  
Db |||||  
QY 961 TTTTGTGAATTTTCCAGATGACCCCAATGAGAGTGTGATGATGATGATGATGATGATGATG 1020  
Db |||||  
QY 961 TTTTGTGAATTTTCCAGATGACCCCAATGAGAGTGTGATGATGATGATGATGATGATGATG 1020  
Db |||||  
QY 1021 TTGTTGTGCGCCAGAGAGAGACTGAAGTTTGAAGTCTTTTGTGCTGCTGCTGCTGCTGCTG 1080  
Db |||||  
QY 1021 TTGTTGTGCGCCAGAGAGAGACTGAAGTTTGAAGTCTTTTGTGCTGCTGCTGCTGCTGCTG 1080  
Db |||||  
QY 1081 TCTAATATTGACTGGAAACAACTTCGTAACCTCTCCCTCCCTCCCTCCCTCCCTCCCTCAAG 1140  
Db |||||  
QY 1081 TCTAATATTGACTGGAAACAACTTCGTAACCTCTCCCTCCCTCCCTCCCTCCCTCAAG 1140  
Db |||||  
QY 1141 TCTGACGATGACACTCCAAATTTTGTGAAACAGAGAGAAATTCGTGGGTTTTCATCCTCT 1200  
Db |||||  
QY 1141 TCTGACGATGACACTCCAAATTTTGTGAAACAGAGAGAAATTCGTGGGTTTTCATCCTCT 1200  
Db |||||  
QY 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGTTTTCG 1260  
Db |||||  
QY 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGTTTTCG 1260  
Db |||||  
QY 1261 TACAGCAGGCACTGGGCAATTTCTGTTAGATCTGAGTCTCTGTTGCTGCGGTTTTCG 1320  
Db |||||  
QY 1261 TACAGCAGGCACTGGGCAATTTCTGTTAGATCTGAGTCTCTGTTGCTGCGGTTTTCG 1320  
Db |||||  
QY 1321 CCGTCCAGACTGCTGCTTCAAGAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1380  
Db |||||  
QY 1321 CCGTCCAGACTGCTGCTTCAAGAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1380  
Db |||||  
QY 1381 TCTCAGACAAAGTGTCAACAGATGGAGAGAAATGACCCGTTTACATCGGAGAGTCA 1440  
Db |||||  
QY 1381 TCTCAGACAAAGTGTCAACAGATGGAGAGAAATGACCCGTTTACATCGGAGAGTCA 1440  
Db |||||  
QY 1441 GAGTGGAGCTGTGCTTGTAGTCAAGAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1500  
Db |||||  
QY 1441 GAGTGGAGCTGTGCTTGTAGTCAAGAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1500  
Db |||||  
QY 1501 TCCCTCTGAGCAGGACCTTGTCTACATCACAAGATGCAAGGCTTCTGCTGCTGCTGCTGCTG 1560  
Db |||||  
QY 1501 TCCCTCTGAGCAGGACCTTGTCTACATCACAAGATGCAAGGCTTCTGCTGCTGCTGCTGCTG 1560  
Db |||||  
QY 1561 TTGAGCAAGCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
Db |||||  
QY 1561 TTGAGCAAGCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
Db |||||  
QY 1621 GATATCAGAGCAGAGCCTGGAAGCTTCCAAAGATCAAGAGAGAGGATGAGGCTGCTGCTG 1680  
Db |||||  
QY 1621 GATATCAGAGCAGAGCCTGGAAGCTTCCAAAGATCAAGAGAGAGGATGAGGCTGCTGCTG 1680  
Db |||||  
QY 1681 GTGGAAGAAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740  
Db |||||  
QY 1681 GTGGAAGAAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740  
Db |||||  
QY 1741 CGGAGTATCTTACGAAATCTGAGCTGAGAGTCTCGGCTTGTGCTGCTGCTGCTGCTGCTGCTG 1800  
Db |||||  
QY 1741 CGGAGTATCTTACGAAATCTGAGCTGAGAGTCTCGGCTTGTGCTGCTGCTGCTGCTGCTGCTG 1800  
Db |||||  
QY 1801 CGGAAAGCAGAGAAATGTGAGCAATAAATGTTGAAGCTTGAAGATCAAGGAGGCTGAA 1860  
Db |||||  
QY 1801 CGGAAAGCAGAGAAATGTGAGCAATAAATGTTGAAGCTTGAAGATCAAGGAGGCTGAA 1860  
Db |||||  
QY 1861 GTGGGAGAAATATCGAAATCGGAAATCGGAAATCGGAAATCGGAAATCGGAAATCGGAAAT 1920  
Db |||||  
QY 1861 GTGGGAGAAATATCGGAAATCGGAAATCGGAAATCGGAAATCGGAAATCGGAAATCGGAAAT 1920  
Db |||||

QY 1921 CTCCAAGAAACTCGAAGGCTG----- 1945  
Db 1921 CTCCAAGAAACTCGAAGGCTGTAAGCCACACGAGGCCACCGAGCTGCTGAG 1980  
QY 1946 -----CAAAAGAGCGAGCCGAGAGGAGCTGGAGAGCTCGAAGCTCGAAGCCGAGAG 1992  
Db 1981 AATATCCGCCAGGCAAAAGGAGCGAGCCGAGAGGAGCTGGAGAGCTCGAAGCCGAGAG 2040  
QY 1993 GATTTCTTGTGAGGCGATCAGAAAGAGCTGGTGGAGCTGAGGAACCGCGCATTTCTCTG 2052  
Db 2041 GATTTCTTGTGAGGCGATCAGAAAGAGCTGGTGGAGCTGAGGAACCGCGCATTTCTCTG 2100  
QY 2053 GAGAAACAAGGTAAAGAGACTAGAGACCCTAGAGCCCTAGAGAAACAGACTGAAGATGAC 2112  
Db 2101 GAGAAACAAGGTAAAGAGACTAGAGACCCTAGAGCCCTAGAGAAACAGACTGAAGATGAC 2160  
QY 2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTCCTGGAGCTCGAAGAG 2172  
Db 2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTCCTGGAGCTCGAAGAG 2220  
QY 2173 AAACATCGGGAGGCCCAAGTCTCAGCCACAGCACTAGAAAGTCACTGAAACAGAAAGAG 2232  
Db 2221 AAACATCGGGAGGCCCAAGTCTCAGCCACAGCACTAGAAAGTCACTGAAACAGAAAGAG 2280  
QY 2233 CAGCACTATGAGGAAAGAAATTAAGTGTGGACAATCAGATATAAGAAAGACTGCTGCTGAC 2292  
Db 2281 CAGCACTATGAGGAAAGAAATTAAGTGTGGACAATCAGATATAAGAAAGACTGCTGCTGAC 2340  
QY 2293 AAGGACACTCGAGAAACATGATGAGACACAGGAGGAGGCCCATGAGAGGCGCAA 2352  
Db 2341 AAGGACACTCGAGAAACATGATGAGACACAGGAGGAGGCCCATGAGAGGCGCAA 2400  
QY 2353 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2412  
Db 2401 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA 2460  
QY 2413 CAGAGATTTGGAACTGCTGAGCCATAAATCTGAGCAAAATAGCAGTCTTTTACC 2472  
Db 2461 CAGAGATTTGGAACTGCTGAGCCATAAATCTGAGCAAAATAGCAGTCTTTTACC 2520  
QY 2473 CAAAGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532  
Db 2521 CAAAGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580  
QY 2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCGAGAACCGGAAACTGGAGAGCAGCTGGAG 2592  
Db 2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCGAGAACCGGAAACTGGAGAGCAGCTGGAG 2640  
QY 2593 AAGATCAGCCACCAAGACACAGTGAACAAGATCGGCTCTGAACTGGAGACAAAGATTG 2652  
Db 2641 AAGATCAGCCACCAAGACACAGTGAACAAGATCGGCTCTGAACTGGAGACAAAGATTG 2700  
QY 2653 CGGGAGGTCAGTCTAGAGCACGAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2712  
Db 2701 CGGGAGGTCAGTCTAGAGCACGAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2760  
QY 2713 CTAAGCTCTCCCTGAGAGCGGAGTACAGTTGACAGCCCTGAGGCTGACGGGG 2772  
Db 2761 CTAAGCTCTCCCTGAGAGCGGAGTACAGTTGACAGCCCTGAGGCTGACGGGG 2820  
QY 2773 GCCTCGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGACCAAGCT 2832  
Db 2821 GCCTCGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGACCAAGCT 2880  
QY 2833 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAAAATTTGATGCT 2892  
Db 2881 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAAAATTTGATGCT 2940  
QY 2893 CTTTCGTAAACAGCTGTACTGTAAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAG 2952  
Db 2941 CTTTCGTAAACAGCTGTACTGTAAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAG 3000  
QY 2953 GACAACGCTGACTCAACCAACAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 3012

Db 3001 GACAACGCTGAACTCAACCAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060  
QY 3013 GGGCCCAACACAGAGATTGTACAACCTGCGAAGTGAAGTGGACCATCTCCCGCGGAGATC 3072  
Db 3061 GGGCCCAACACAGAGATTGTACAACCTGCGAAGTGAAGTGGACCATCTCCCGCGGAGATC 3120  
QY 3073 ACGGAAACGAGAGATGCACTTACAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG 3132  
Db 3121 ACGGAAACGAGAGATGCACTTACAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG 3180  
QY 3133 TGCACCATGCTGAGAGAACAGGTCATGGATTTTGGAGGCCCTAAACGATGAGCTGTAGAA 3192  
Db 3181 TGCACCATGCTGAGAGAACAGGTCATGGATTTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240  
QY 3193 AAGAGCGGAGTGGGAGGCTCGAGAGGCTCTCGGCTGATGAGAAATCCCAAGTTTGG 3252  
Db 3241 AAGAGCGGAGTGGGAGGCTCGAGAGGCTCTCGGCTGATGAGAAATCCCAAGTTTGG 3300  
QY 3253 TGTCCGTTTCCAGAGCTGCGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGAGAGCC 3312  
Db 3301 TGTCCGTTTCCAGAGCTGCGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGAGAGCC 3360  
QY 3313 GATCAGCGGATCCCGAGTCTCCGAGGCTGGAGCTGGAGTGAAGGAGGACCAAGGCT 3372  
Db 3361 GATCAGCGGATCCCGAGTCTCCGAGGCTGGAGCTGGAGTGAAGGAGGACCAAGGCT 3420  
QY 3373 GAGATTTCTCGCTCTGACAGGCTCTCAAAAGACAGAAAGTGAAGGCCGAGAGGCTCTCT 3432  
Db 3421 GAGATTTCTCGCTCTGACAGGCTCTCAAAAGACAGAAAGTGAAGGCCGAGAGGCTCTCT 3480  
QY 3433 GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGATGATGCTCCGAGGCTTA 3492  
Db 3481 GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGATGATGCTCCGAGGCTTA 3540  
QY 3493 CAGCAGAGCTGGAGACTGAACAGAGACTCAAAACAGAGGCTTCTGGAGAGGACCAAGCA 3552  
Db 3541 CAGCAGAGCTGGAGACTGAACAGAGACTCAAAACAGAGGCTTCTGGAGAGGACCAAGCA 3600  
QY 3553 TTAACAGAGCAGATGGAACCTGCAAGAAATCAATTTTCCGTCTGATCTCAAGGACTGCA 3612  
Db 3601 TTAACAGAGCAGATGGAACCTGCAAGAAATCAATTTTCCGTCTGATCTCAAGGACTGCA 3660  
QY 3613 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAAGTGAAGTATCAGCTG 3672  
Db 3661 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAAGTGAAGTATCAGCTG 3720  
QY 3673 GAAACCATTCAGGTTCTTATTTCTATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3732  
Db 3721 GAAACCATTCAGGTTCTTATTTCTATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3780  
QY 3733 CAAACCAACTCATTTGATTTTCTCAGAGCCAAATGACCAACCTGCTTAAAGAAAGAA 3792  
Db 3781 CAAACCAACTCATTTGATTTTCTCAGAGCCAAATGACCAACCTGCTTAAAGAAAGAA 3838  
QY 3793 GGTATTATTAGTCGACGAAAGAGGACCTGCTTTTACCACACAGGTTCTCTCAGTAC 3852  
Db 3839 -----AGGTTCTCTCAGTAC 3855  
QY 3853 AATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGTGTGAGAGCTAGAGAGGAC 3912  
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGTGTGAGAGCTAGAGAGGAC 3915  
QY 3913 CTTTCAGAGACCCGATCAGCTCCGCTCCGCGGAGGAGGCTGCCACCGCAAGCA 3972  
Db 3916 CTTTCAGAGACCCGATCAGCTCCGCTCCGCGGAGGAGGCTGCCACCGCAAGCA 3975  
QY 3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGAGATGCGCATGCTCCGCC 4032  
Db 3976 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGAGATGCGCATGCTCCGCC 4035  
QY 4033 ATCTGGGCTGCGCAGAGCACCGCCAGTGCCTGCTGCTGGCCCGCCCTCAGC 4092

Db 4036 ATCGTGGGTGCGCAGAGCACCAGCCAGTGCATGAGCCTGTGCGCCCGCCATCCAGC 4095  
Qy 4093 CGCAGAAAGAGTCTTCAATCCAGAGGAATTTAGTCGGCGTCTTTAAGGAACCCATGCAC 4152  
Db 4096 CGCAGAAAGAGTCTTCAATCCAGAGGAATTTAGTCGGCGTCTTTAAGGAACCCATGCAC 4155  
Qy 4153 CACAATATCTCTCACCAGATTCAACGTAGGACTGAACATGAGCCAGCCAAAGTGTGCTGTG 4212  
Db 4156 CACAATATCTCTCACCAGATTCAACGTAGGACTGAACATGAGCCAGCCAAAGTGTGCTGTG 4215  
Qy 4213 TGTCTGATACCGTGCATTTGGACGCCAGGCATCCAAATGTCGAAATGTCAGGTGATG 4272  
Db 4216 TGTCTGATACCGTGCATTTGGACGCCAGGCATCCAAATGTCGAAATGTCAGGTGATG 4275  
Qy 4273 TGTACACCCAGTGTCCAGTGTCTGCGAGCCACCTGGCGTGTGCTGTGAATATGCC 4332  
Db 4276 TGTACACCCAGTGTCCAGTGTCTGCGAGCCACCTGGCGTGTGCTGTGAATATGCC 4335  
Qy 4333 ACACATTCACCGAGGCTTCTGCGGTGACAAATGAATCCCCAGGTCTCCAGACCAAG 4392  
Db 4336 ACACATTCACCGAGGCTTCTGCGGTGACAAATGAATCCCCAGGTCTCCAGACCAAG 4395  
Qy 4393 GAGCCAGAGCAGCTTGACCTTGAAGGTGATGAGTGCACAGGAATCAACAACGA 4452  
Db 4396 GAGCCAGAGCAGCTTGACCTTGAAGGTGATGAGTGCACAGGAATCAACAACGA 4455  
Qy 4453 GGACAGCAAGCTGGGACAGGAATGATTTCTGAGGGATCAAAAGTCTCTCATTTAT 4512  
Db 4456 GGACAGCAAGCTGGGACAGGAATGATTTCTGAGGGATCAAAAGTCTCTCATTTAT 4515  
Qy 4513 GACAAATGAAGCCAGAGAGTGTGACAGAGCCGGTGGAAAGATTTGAGTGTGCTTCCC 4572  
Db 4516 GACAAATGAAGCCAGAGAGTGTGACAGAGCCGGTGGAAAGATTTGAGTGTGCTTCCC 4575  
Qy 4573 GACGGGATGTATCTATCATGTGCGGTGGTGTCTCCGAACTGCAATACAGCCAAA 4632  
Db 4576 GACGGGATGTATCTATCATGTGCGGTGGTGTCTCCGAACTGCAATACAGCCAAA 4635  
Qy 4633 GCA----- 4635  
Db 4636 GCAGATGCCATACATACATGGAATCTCACCCGACACACCTGCTGGCCCGG 4695  
Qy 4636 ----- 4635  
Db 4696 AGAACCTCTACTTGTAGTCCAGTCCAGTCCCTGTGACAAACAGCGTGGGTACCCGCTTA 4755  
Qy 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4662  
Db 4756 GAATCAGTTGTGCGAGTGGAGAGTTTCTAGGAAAAGCAGAGCTGATGCTAACTG 4815  
Qy 4663 CTTGGAATCTCCCTGTGAAACTGGAAGTGTATGACCTGTAGACATGATGCAAGTGTG 4722  
Db 4816 CTTGGAATCTCCCTGTGAAACTGGAAGTGTATGACCTGTAGACATGATGCAAGTGTG 4875  
Qy 4723 CCCTTCAGTGACAGAGTGTGTTGGTGGCACCGAGGAAGGCTCTAGCCCTGATGTC 4782  
Db 4876 CCCTTCAGTGACAGAGTGTGTTGGTGGCACCGAGGAAGGCTCTAGCCCTGATGTC 4935  
Qy 4783 TTGAAAAATCCCTTAACCATGTCACAGAAATGGAGCAGTCTTCCAAATTTATATTATC 4842  
Db 4936 TTGAAAAATCCCTTAACCATGTCACAGAAATGGAGCAGTCTTCCAAATTTATATTATC 4995  
Qy 4843 AAGGACCTGGAGAGCTACTCATGTATGACAGAGAGCGGCACTGTCTTGTGGAC 4902  
Db 4996 AAGGACCTGGAGAGCTACTCATGTATGACAGAGAGCGGCACTGTCTTGTGGAC 5055  
Qy 4903 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 4962  
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 5115  
Qy 4963 CCCACATTTTGAAGCTGTCAAGGCTGCCACTTTGTTGGGAGGCAAGATTCAGAAC 5022  
Db 5116 CCCACATTTTGAAGCTGTCAAGGCTGCCACTTTGTTGGGAGGCAAGATTCAGAAC 5175

Qy 5023 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTGTCATTTCTCCGTACAAACGAAAC 5082  
Db 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAAAGTGTGTCATTTCTCCGTACAAACGAAAC 5235  
Qy 5083 CTCACAAATACTCATCCGGAAGAGATGAGACCTCAGAGCCCTGAGCTGTATCCAC 5142  
Db 5236 CTCACAAATACTCATCCGGAAGAGATGAGACCTCAGAGCCCTGAGCTGTATCCAC 5295  
Qy 5143 TTCACAAATACAGTATCTCTCAITGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5202  
Db 5296 TTCACAAATACAGTATCTCTCAITGGAAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5355  
Qy 5203 TACAGCTCGAGGAATTCCTGGATAAGAAATGACCAATTTCTTGGCACTGTGTGTTGCC 5262  
Db 5356 TACAGCTCGAGGAATTCCTGGATAAGAAATGACCAATTTCTTGGCACTGTGTGTTGCC 5415  
Qy 5263 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGGTGAACAGCGAGCGAGCGAGAG 5322  
Db 5416 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGAGGTGAACAGCGAGCGAGCGAGAG 5475  
Qy 5323 GAGTACTTGTGTGTTTCCACGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5382  
Db 5476 GAGTACTTGTGTGTTTCCACGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5535  
Qy 5383 CGCACAGACGATCTCAAGTGGAGTGTGCTTACCTTTGGCCTTTGGCTACAGAGAACCTAT 5442  
Db 5536 CGCACAGACGATCTCAAGTGGAGTGTGCTTACCTTTGGCCTTTGGCTACAGAGAACCTAT 5595  
Qy 5443 CTGTTTGTGACCCACTTCAACTCACTCAAGTAAATTTAGAGTAAATTTAGGGTCAATTTGCTGC 5622  
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCAAGTAAATTTAGAGTAAATTTAGGGTCAATTTGCTGC 5775  
Qy 5503 GGGACCTCTGCCAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCAAT 5562  
Db 5556 GGGACCTCTGCCAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCAAT 5715  
Qy 5563 TCCTTCAGAGCGATTTACTTTGGGTCTCTATACAGGATAAATTTAGGGTCAATTTGCTGC 5622  
Db 5716 TCCTTCAGAGCGATTTACTTTGGGTCTCTATACAGGATAAATTTAGGGTCAATTTGCTGC 5775  
Qy 5623 AAGGAAACCTCTGAGAGAGTCCGGCACTGAACACCGCGGCGCTCAGCTCCCGC 5682  
Db 5776 AAGGAAACCTCTGAGAGAGTCCGGCACTGAACACCGCGGCGCTCAGCTCCCGC 5835  
Qy 5683 AGCAGCCCCAACAGCGAGGCCACCCACGTACAAACAGACACATCAACAGCGCTGGCC 5742  
Db 5836 AGCAGCCCCAACAGCGAGGCCACCCACGTACAAACAGACACATCAACAGCGCTGGCC 5895  
Qy 5743 TCAGCCAGCGCGCGCGCGAGGCCACCGACCCAGCGAGAGCAACACCCACCGC 5802  
Db 5896 TCAGCCAGCGCGCGCGCGAGGCCACCGACCCAGCGAGCAACAGCACACCCACCGC 5955  
Qy 5803 TACCGAGGGGGGAGCCGAGCTGCGCAGGACAAAGTCTCTGCGCGCGCGCTGGAGCGA 5862  
Db 5956 TACCGAGGGGGGAGCCGAGCTGCGCAGGACAAAGTCTCTGCGCGCGCGCTGGAGCGA 6015  
Qy 5863 GAGAAGTCCCGCGCGGATGCTCAGCAACCGAGAGAGCGGTCCCGCGGAGCGTGT 5922  
Db 6016 GAGAAGTCCCGCGCGGATGCTCAGCAACCGAGAGAGCGGTCCCGCGGAGCGTGT 6075  
Qy 5923 GAGACAGACAGAGGCGCGCTGCTGGGAGCGTGGAGACCGCTGTGTCGAGGTTG 5982  
Db 6076 GAGACAGACAGAGGCGCGCTGCTGGGAGCGTGGAGACCGCTGTGTCGAGGTTG 6135  
Qy 5983 AACAGGAAAGAGGAGGAGTGTCTCAGTAAATTTTACGGTTTTCACGGTTTAACTGTCACTTAT 6042  
Db 6136 AACAGGTTCTGGACAGTCTTTCAGTAAATTTCTCAGGAGAAACCAACTCTCTCATCT 6195

DEFINITION Sequence 8 from Patent WO03004523.  
ACCESSION AX671112  
VERSION AX671112.1 GI:29329572  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1.  
AUTHORS Zhu, Z.  
TITLES Regulation of human citron rho/rac-interacting kinase  
JOURNAL Patent: WO 03004523-A 8 16-JAN-2003;  
Bayer Aktiengesellschaft (DE)  
FEATURES  
Location/Qualifiers  
1..6156  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 92.0%; Score 5666.6; DB 6; Length 6156;  
Best Local Similarity 95.9%; Pred. No. 0;  
Matches 5945; Conservative 0; Mismatches 9; Indels 243; Gaps 3;  
QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGGTGTGCTGAACCCATT 60  
DB 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGGTGTGCTGAACCCATT 60  
QY 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCCAGGGGAACACCCCTTTATGACTCAA 120  
DB 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAACACCCCTTTATGACTCAA 120  
QY 121 CAGCAGATGTCCTCTTTCCGAGAGGATATATAGATGCCCTTTGTCTCTTTGAA 180  
DB 121 CAGCAGATGTCCTCTTTCCGAGAGGATATATAGATGCCCTTTGTCTCTTTGAA 180  
QY 181 GAATCAGTACGCTGCTCTGATGAAGATTAGACGCTGAGCACTTTGTCGGGAAGTAT 240  
DB 181 GAATCAGTACGCTGCTCTGATGAAGATTAGACGCTGAGCACTTTGTCGGGAAGTAT 240  
QY 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 300  
DB 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 300  
QY 301 AGTCTTGTAGTTGTGGTCACTTTGCTGAAGTGCAGGTGTTAAGAGAGAACACCGGG 360  
DB 301 AGTCTTGTAGTTGTGGTCACTTTGCTGAAGTGCAGGTGTTAAGAGAGAACACCGGG 360  
QY 361 GACATCTATGCTATGAAGTGTGAAGAGAGAGGCTTTATTGGCCCGAGGACAGTTTCA 420  
DB 361 GACATCTATGCTATGAAGTGTGAAGAGAGAGGCTTTATTGGCCCGAGGACAGTTTCA 420  
QY 421 TTTTGTAGGAAGAGCGGAACATATATCTCGAAGCACAAGCCCGTGGATCCCCAATTA 480  
DB 421 TTTTGTAGGAAGAGCGGAACATATATCTCGAAGCACAAGCCCGTGGATCCCCAATTA 480  
QY 481 CAGTATGCTTTTTCAGGACAAAATCACCTTTATCTGATGGAGATATCAGCTGGAGGG 540  
DB 481 CAGTATGCTTTTTCAGGACAAAATCACCTTTATCTGATGGAGATATCAGCTGGAGGG 540  
QY 541 GACTTGTCTCACTTTTGAATAGATATGAGGACCACTAGATGAAAACCTGATACAGTTT 600  
DB 541 GACTTGTCTCACTTTTGAATAGATATGAGGACCACTAGATGAAAACCTGATACAGTTT 600  
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACACGCTTCTATCTGATGGATACGTGATCGA 660  
DB 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACACGCTTCTATCTGATGGATACGTGATCGA 660  
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCAGAGGACATCAAGCTGGTGAATTT 720  
DB 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCAGAGGACATCAAGCTGGTGAATTT 720  
QY 721 GGATCTGCCCGGAAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780

721 GGATCTGCCCGGAAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780  
QY 781 CCAGATTACATGCTCCTGAACTGCTGACTGTGATGAACGGGATGAAAGACACCTAC 840  
DB 781 CCAGATTACATGCTCCTGAACTGCTGACTGTGATGAACGGGATGAAAGACACCTAC 840  
QY 841 GGCTGACATGCTGACTGGTGGTCACTGGGCGTGAATTCCTATGAGATGATTTATGGGAGA 900  
DB 841 GGCTGACATGCTGACTGGTGGTCACTGGGCGTGAATTCCTATGAGATGATTTATGGGAGA 900  
QY 901 TCCCTCTTCCAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTAATGAATTTCCAGCGG 960  
DB 901 TCCCTCTTCCAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTAATGAATTTCCAGCGG 960  
QY 961 TTTTGAATTTTCCAGATGATCCCAAGTCCAGAGTGAATTTCTTGAATTCGATTCAGATC 1020  
DB 961 TTTTGAATTTTCCAGATGATCCCAAGTCCAGAGTGAATTTCTTGAATTCGATTCAGATC 1020  
QY 1021 TTTTGTGCCGCCAGAAAGAGAGACTGAATTTGAAGGTCTTTGTGCCATCTTTCTTTTC 1080  
DB 1021 TTTTGTGCCGCCAGAAAGAGAGACTGAATTTGAAGGTCTTTGTGCCATCTTTCTTTTC 1080  
QY 1081 TCTAAATTTGACTGGACACATTCGTAATCTCTCCCTCCCTTCCTGTTCCACCTCAAG 1140  
DB 1081 TCTAAATTTGACTGGACACATTCGTAATCTCTCTCCCTCTTCTGTTCCACCTCAAG 1140  
QY 1141 TCTGACATGACACCTTCCAAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTCATCTCT 1200  
DB 1141 TCTGACATGACACCTTCCAAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTTCATCTCT 1200  
QY 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTGGAAGAACTGCGGTTTGTGGGTTTTTCG 1260  
DB 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTGGAAGAACTGCGGTTTGTGGGTTTTTCG 1260  
QY 1261 TACAGCAAGGCTGCGGATTTCTTGGTAGATCTGAGTCTGTTGTCGGGCTTGGACTCC 1320  
DB 1261 TACAGCAAGGCTGCGGATTTCTTGGTAGATCTGAGTCTGTTGTCGGGCTTGGACTCC 1320  
QY 1321 CCGTGCAAGCTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1380  
DB 1321 CCGTGCAAGCTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTTACAAGAC 1380  
QY 1381 TCTCAGCAAGTGTCAAAAGTGGACGAGAAATGACCCGGTTTACATCGAGAGTGTCA 1440  
DB 1381 TCTCAGCAAGTGTCAAAAGTGGACGAGAAATGACCCGGTTTACATCGAGAGTGTCA 1440  
QY 1441 GAGGTGAGGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
DB 1441 GAGGTGAGGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
QY 1501 TCCCTCTTGGAGGAGGACCTTGTCTACCTACATCACAGAAATGCAGTAGCTTAAAGCGAAGT 1560  
DB 1501 TCCCTCTTGGAGGAGGACCTTGTCTACCTACATCACAGAAATGCAGTAGCTTAAAGCGAAGT 1560  
QY 1561 TTGGAGCAAGCAGGATGGAGGTGTCCAGAGGAGATGACAAAGCACTGCGAGCTTCTCCAT 1620  
DB 1561 TTGGAGCAAGCAGGATGGAGGTGTCCAGAGGAGATGACAAAGCACTGCGAGCTTCTCCAT 1620  
QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGAGTACCGAGCTCAA 1680  
DB 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGAGTACCGAGCTCAA 1680  
QY 1681 GTGGAAGAAATGAGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740  
DB 1681 GTGGAAGAAATGAGTTGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740  
QY 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGAGTCTGGGCTTGTCTGCTGAAGAAATTCAG 1800  
DB 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGAGTCTGGGCTTGTCTGCTGAAGAAATTCAG 1800  
QY 1801 CGGAAACGACAGATGTCAAGTAAACTGTTGAAGCTAAGGATCAAGGAGAGCTGAA 1860





Db	6136	AACAAGGTGAGGCAGCA	6152	Db	595	GACTTGTGTCACTTTTGAATAGATATGAGGACCGACGTTAGATGAAAACCTGTATACAGTTT	654	
RESULT 5				Qy	601	TACCTAGCTGAGCTGAATTTTGGCTGTGTTCAAGCGTTTCATCTGATGGGATACGTGCATCGA	660	
AX504254				Db	655	TACCTAGCTGAGCTGAATTTTGGCTGTGTTCAAGCGTTTCATCTGATGGGATACGTGCATCGA	714	
LOCUS	AX504254	6298 bp	DNA	linear	PAT 27-SEP-2002			
DEFINITION	Sequence 43 from Patent WO0233099.			Qy	661	GACATCAAGCTGAGAACATTCCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT	720	
ACCESSION	AX504254			Db	715	GACATCAAGCTGAGAACATTCCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT	774	
VERSION	AX504254.1	GI:23386094		Qy	721	GGATCTCCGCGAAAATGAATTCAAAACAAGATGGTGAATGCCAAATCCCGATTGGGACC	780	
KEYWORDS	Homo sapiens (human)			Db	775	GGATCTCCGCGAAAATGAATTCAAAACAAGATGGTGAATGCCAAATCCCGATTGGGACC	834	
SOURCE	Homo sapiens			Qy	781	CCAGATTACATGGCTCTCGAAGTGTGCTGACTGTGATGAACGGGGATGAAAAGGCACCTAC	840	
ORGANISM	Homo sapiens			Db	835	CCAGATTACATGGCTCTCGAAGTGTGCTGACTGTGATGAACGGGGATGAAAAGGCACCTAC	894	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			Qy	841	GGCTGACAGTGTGACTGGTGTGCTGAGTGGGCGTGATTGCCTATCAGATGATGATTTATGGGAGA	900	
AUTHORS	Gururajan, R., Baughn, M.R., Wallia, N.K., Elliott, V.S., Xu, Y., Arvizu, C., Yao, M.G., Ramkumar, J., Ding, L., Tang, Y.T., Hafalia, A.J., Nguyen, D.B., Gandhi, A.R., Lu, Y., Yue, H., Burford, N., Bandman, O., Tribouley, C.M., Lal, P.G., Recipon, S.A., Lu, D.A., Borowsky, M.L., Thornton, M., Swannaker, A., Thangavelu, K., Khan, F.A. and Ison, C.H. Human kinases			Db	895	GGCTGACAGTGTGACTGGTGTGCTGAGTGGGCGTGATTGCCTATCAGATGATGATTTATGGGAGA	954	
TITLE	Patent: WO 0233099-A 43 25-APR-2002;			Qy	901	TCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAACATTATGATGATTTTCCAGCGG	960	
JOURNAL	Incyte Genomics, Inc. (US)			Db	955	TCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAACATTATGATGATTTTCCAGCGG	1014	
FEATURES	Location/Qualifiers			Qy	961	TTTTTGAATTTCCAGATGACCCCAAGTGACAGTGACCTTCTTGATCTGATTTCAAAGC	1020	
source	1..6298	organism="Homo sapiens"		Db	1015	TTTTTGAATTTCCAGATGACCCCAAGTGACAGTGACCTTCTTGATCTGATTTCAAAGC	1074	
ORIGIN		/mol_type="unassigned DNA"		Qy	1021	TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCTGCCATCCTTTCTTC	1080	
Query Match	92.0%;	Score 5666.4;	DB 6;	Length 6298;	Db	1075	TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCTGCCATCCTTTCTTC	1134
Best Local Similarity	95.5%;	Pred. No. 0;			Qy	1081	TCTAAAATTGACTGGAAACACACTTCTGTAACCTCTCCCTCCCTTCTGCCACCTCAAG	1140
Matches 5961; Conservative	0;	Mismatches 36;	Indels 243;	Gaps 3;	Db	1135	TCTAAAATTGACTGGAAACACACTTCTGTAACCTCTCTCCCTCCCTTCTGCCACCTCAAG	1194
Qy	1	ATGTTGAAGTTCAATATATGAGCGCGGAATCCCTTTGGATGCTGGTCTGCTGAACCCAT	60	Qy	1141	TCTGACCATGACACTCCAAATTTTGATGAACACAGAGAAGAAATTCGTGGGTTTCATCCTCT	1200	
Db	55	ATGTTGAAGTTCAATATATGAGCGCGGAATCCCTTTGGATGCTGGTCTGCTGAACCCAT	114	Db	1195	TCTGACCATGACACTCCAAATTTTGATGAACACAGAGAAGAAATTCGTGGGTTTCATCCTCT	1254	
Qy	61	GCAGCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCAA	120	Qy	1201	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCCGTTTGTGGGGTTTTCG	1260	
Db	115	GCAGCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCAA	174	Db	1255	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG	1314	
Qy	121	CACGAGATGTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTCTCTTTGAA	180	Qy	1261	TACAGCAAGCAGCTGGGGATTTCTGGTGTAGATCTGAGTCTGTTGTGTCGGGTCTGGACTCC	1320	
Db	175	CACGAGATGTCCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTCTCTTTGAA	234	Db	1315	TACAGCAAGCAGCTGGGGATTTCTGGTGTAGATCTGAGTCTGTTGTGTCGGGTCTGGACTCC	1374	
Qy	181	GAATGAGTCAGCTGCTCTGATGAAGATTAAGCAGCTGAGCAAATTTGTCGGGAAGTAT	240	Qy	1321	CCTGCCAAGACTAGCTCCATGGAAAAGAAATCTTCTATCAAAAAGCAAGAGCTACAGAC	1380	
Db	235	GAATGAGTCAGCTGCTCTGATGAAGATTAAGCAGCTGAGCAAATTTGTCGGGAAGTAT	294	Db	1375	CCTGCCAAGACTAGCTCCATGGAAAAGAAATCTTCTATCAAAAAGCAAGAGCTACAGAC	1434	
Qy	241	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGAGACTTCGAAGTCAGA	300	Qy	1381	TCTCAGGACAGTGTCAAGATGGAGCAAGAAATGACCCGTTTACATCCGAGAGTGTCA	1440	
Db	295	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGAGACTTCGAAGTCAGA	354	Db	1435	TCTCAGGACAGTGTCAAGATGGAGCAGAAATGACCCGGTTTACATCCGAGAGTGTCA	1494	
Qy	301	AGTCTTGTAGTTTGTGTCATTTGCTGAAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG	360	Qy	1441	GAGTGGAGGCTGTGCTTAGTTCAGAAGGAGGTGGAGCTGAAGCGCTCTGAGACTCAGAGA	1500	
Db	355	AGTCTTGTAGTTTGTGTCATTTGCTGAAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG	414	Db	1495	GAGTGGAGGCTGTGCTTAGTTCAGAAGGAGGTGGAGCTGAAGGCTCTCAGACTCAGAGA	1554	
Qy	361	GACATCTAGCTATGAAGTGTAGAGAGAGGCTTTATGGCCAGGAGCAGGTTTCA	420	Qy	1501	TCCCTCTCGAGCAGGACCTTGCTACCTACATCACAGAAATGCGATAGCTTTAAAGCGAAGT	1560	
Db	415	GACATCTAGCTATGAAGTGTAGAGAGAGGCTTTATGGCCAGGAGCAGGTTTCA	474	Db	1555	TCCCTCTCGAGCAGGACCTTGCTACCTACATCACAGAAATGCGATAGCTTTAAAGCGAAGT	1614	
Qy	421	TTTTTTGAGGAGAGCGGAAATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA	480	Qy	1561	TTGAGCAAGCACCGATGGAGGTGTCCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT	1620	
Db	475	TTTTTTGAGGAGAGCGGAAATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA	534	Db	1615	TTGAGCAAGCACCGATGGAGGTGTCCAGGAGGATGACAAAGCACTGCAGCTTCTCCAT	1674	
Qy	481	CAGTATGCCCTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGCTGGAGGG	540	Qy	1621	GATATCAGAGCAGAGCCCGAGCTCCAGAAATCAAGAGCAGGAGTACCGAGCTCAA	1680	
Db	535	CAGTATGCCCTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGCTGGAGGG	594	Db	1675	GATATCAGAGCAGAGCCCGAGACTCCAGAAAATAAGAGCAGGAGTACCGAGCTCAA	1734	

QY	1681	GTGGAGAAATGAGTTGATGATGAATCAGTTGGAGAGAGATCTTGCTCAGCAAGAAGA	1740
Db	1735	GTGGAAGAAATGAGTTGATGATGAATCAGTTGGAGAGAGATCTTGCTCAGCAAGAAGA	1794
QY	1741	CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATTCAG	1800
Db	1795	CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATTCAG	1854
QY	1801	CGGAAGCCACAGATGTCAGCATAACTGTTGAAGGCTAAGGATCAAGGGAAGCTGAA	1860
Db	1855	CGGAAGCCACAGATGTCAGCATAACTGTTGAAGGCTAAGGATCAAGGGAAGCTGAA	1914
QY	1861	GTGGAGAAATATCGGAAAATCGGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1920
Db	1915	GTGGAGAAATATCGGAAAATCGGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1974
QY	1921	CTCCAAAGAAACTGGAGAAAGCTG	1945
Db	1975	CTCCAAAGAAACTGGAGAAAGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGCTGCAG	2034
QY	1946	-----CAAAGGAGCGAGCGAGAGGGAGCTGGAGAGCTCGAAGCCGAGAG	1992
Db	2035	AAATATCCGCGAGCAAGAGGAGCGCCGAGAGGGAGCTGGAGAGCTCGAAGCCGAGAG	2094
QY	1993	GATTCTTCTGAAGGCATCAAGAAAGAGCTGGTGGAGCTGAGGAAACGCGCCCATTTCTCTG	2052
Db	2095	GATTCTTCTGAAGGCATCAAGAAAGAGCTGGTGGAGCTGAGGAAACGCGCCCATTTCTCTG	2154
QY	2053	GAGACAGAGTAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGACTGAGAGATGAC	2112
Db	2155	GAGACAGAGTAAGAGACTAGAGACCATGGAGCGTAGAGAAACAGACTGAGAGATGAC	2214
QY	2113	ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAAGAG	2172
Db	2215	ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTTCTGGAGCTCGAAGAG	2274
QY	2173	AAACATCGGAGGCCAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTGAAACAGAGAGAG	2232
Db	2275	AAACATCGGAGGCCAAGTCTCAGCCAGCAGCTAGAGAGTGCACCTGAAACAGAGAGAG	2334
QY	2233	CAGCACTATGAGGAAAGATTAAGTTGTTGGAACATCAGATTAAGAAAGACTGGCTGAC	2292
Db	2335	CAGCACTATGAGGAAAGATTAAGTTGTTGGAACATCAGATTAAGAAAGACTGGCTGAC	2394
QY	2293	AAGGAGACACTGGAGAACATGATGAGAGACACGAGGAGGAGGCCCATGAGAAGGCCAAA	2352
Db	2395	AAGGAGACACTGGAGAACATGATGAGAGACACGAGGAGGAGGCCCATGAGAAGGCCAAA	2454
QY	2353	ATTCTCAGGACAGAGGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA	2412
Db	2455	ATTCTCAGGACAGAGGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTGGAA	2514
QY	2413	CAGAGGATTGTGGAACCTGTCTGAAGCCCAATAACTTGCAGCAAAATAGCAGTCTTTTACC	2472
Db	2515	CAGAGGATTGTGGAACCTGTCTGAAGCCCAATAACTTGCAGCAAAATAGCAGTCTTTTACC	2574
QY	2473	CAAAGGACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC	2532
Db	2575	CAAAGGACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC	2634
QY	2533	CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCCGAAACTGGAGGAGCAGCTGGAG	2592
Db	2635	CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCCGAAACTGGAGGAGCAGCTGGAG	2694
QY	2593	AAGATCAGCCCAACAGACCAAGTGAAGAAATCGGCTGCTGGAACCTGGAGACAAGATTG	2652
Db	2695	AAGATCAGCCCAACAGACCAAGTGAAGAAATCGGCTGCTGGAACCTGGAGACAAGATTG	2754
QY	2653	CGGAGGCTCAGTCTAGAGACCAAGAGAGCAGAAACTGGAGCTCAGCGCCAGCTCAGAGAG	2712
Db	2755	CGGAGGCTCAGTCTAGAGACCAAGAGAGCAGAAACTGGAGCTCAGCGCCAGCTCAGAGAG	2814

QY	2713	CTACAGCTTCTCCCTGCAGGAGCGCGAGTCAAGTTGACAGCCCTGCAGGCTGCACGGGCG	2772
Db	2815	CTACAGCTTCTCCCTGCAGGAGCGCGAGTCAAGTTGACAGCCCTGCAGGCTGCACGGGCG	2874
QY	2773	GCCCTGAGAGAGCCAGCTTCCGCGAGCGAAGACAGAGCTGGAAGAGACCAAGAGAGAGCT	2832
Db	2875	GCCCTGAGAGAGCCAGCTTCCGCGAGCGAAGACAGAGCTGGAAGAGACCAAGAGAGAGCT	2934
QY	2833	GAAGAGAGATCAGGAGCACTCAAGGCATAGAGATGAAATCCAGCGCAAAATTTGATGCT	2892
Db	2935	GAAGAGAGATCAGGAGCACTCAAGGCATAGAGATGAAATCCAGCGCAAAATTTGATGCT	2994
QY	2893	CTTCCTAACAGCTGCTACTGTAATCACAGACTCGAGGAGCAGCTTAAACAGCTGACCGAG	2952
Db	2995	CTTCCTAACAGCTGCTACTGTAATCACAGACTCGAGGAGCAGCTTAAACAGCTGACCGAG	3054
QY	2953	GACAACTGTAACCTCAACAAACCAAACTTCTAATTGTTCCAAAACAACTCGATAGGCTTCT	3012
Db	3055	GACAACTGTAACCTCAACAAACCAAACTTCTAATTGTTCCAAAACAACTCGATAGGCTTCT	3114
QY	3013	GGGCCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACCATCTCCGCGGGAGATC	3072
Db	3115	GGGCCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACCATCTCCGCGGGAGATC	3174
QY	3073	ACGGAACGAGAGATGACAGCTTACAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG	3132
Db	3175	ACGGAACGAGAGATGACAGCTTACAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG	3234
QY	3133	TGCACCATGCTGGAGGAAACAGGTCATGGATTGGAGGCCCTTAAACGATAGCTGCTAGAA	3192
Db	3235	TGCACCATGCTGGAGGAAACAGGTCATGGATTGGAGGCCCTTAAACGATAGCTGCTAGAA	3294
QY	3193	AAAGAGCGGAGTGGGAGGCTCGAGGAGCGTCTCGGCTGATGAGAAATCCAGTTTGAG	3252
Db	3295	AAAGAGCGGAGTGGGAGGCTCGAGGAGCGTCTCGGCTGATGAGAAATCCAGTTTGAG	3354
QY	3253	TGTCGGGTTTCGAGAGCTGACAGAAATGCTGGAACCCAGAGAAAACAGAGCGGAGAGCC	3312
Db	3355	TGTCGGGTTTCGAGAGCTGACAGAAATGCTGGAACCCAGAGAAAACAGAGCGGAGAGCC	3414
QY	3313	GATCAGCGGATCAGGAGTCTCCGAGGCTGCGAGCTGCGAGTGAAGGAGCAACAGGCT	3372
Db	3415	GATCAGCGGATCAGGAGTCTCCGAGGCTGCGAGTGAAGGAGCAACAGGCT	3474
QY	3373	GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGACAGAAAGCTGAAGGCCGAGAGCCCTCTCT	3432
Db	3475	GAGATTCTCGCTCTGCAGCAGGCTCTCAAAGACAGAAAGCTGAAGGCCGAGAGCCCTCTCT	3534
QY	3433	GACAAGCTCAATGACCTGGAGAGAGATGCTATGCTTGAATGAATGATGCCCCGAAGCTTA	3492
Db	3535	GACAAGCTCAATGACCTGGAGAGAGATGCTATGCTTGAATGAATGATGCCCCGAAGCTTA	3594
QY	3493	CAGCAGAAAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTCTGGAAGCAGCAAGCCAA	3552
Db	3595	CAGCAGAAAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTCTGGAAGCAGCAAGCCAA	3654
QY	3553	TTACAGCAGCAGATGAGCTGCAAGAAAATCAATTTTCCGTTCTGACTCAAGGACTGCA	3612
Db	3655	TTACAGCAGCAGATGAGCTGCAAGAAAATCAATTTTCCGTTCTGACTCAAGGACTGCA	3714
QY	3613	GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAACCTGGAGTATCAGCTG	3672
Db	3715	GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAACCTGGAGTATCAGCTG	3774
QY	3673	GAATACTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGCACTATTTCTCAA	3732
Db	3775	GAATACTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGCACTATTTCTCAA	3834
QY	3733	CAAAACCAACTCATGATTTTCTGGAAGCCAAAATGGAACCAACCTGCTAAAAAGAAAAAG	3792
Db	3835	CAAAACCAACTCATGATTTTCTGGAAGCCAAAATGGAACCAACCTGCTAAAAAGAAAA	3892
QY	3793	GGTTTATTTAGTCGACGGAAGAGGAGCCCTGCTTTTACCACACACAGGTTCTCTCTCAGTAC	3852



QY	5863	GAGAAAGTCCCGCGCGGATGCTCAGACGCGGAGAGAGCGGTCCCGCGGAGGCTGTTT	5922
Db	6070	GAGAAAGTCCCGCGCGGATGCTCAGACGCGGAGAGAGCGGTCCCGCGGAGGCTGTTT	6129
QY	5923	GAAGACAGCAGCAGGCGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCACAGGTG	5982
Db	6130	GAAGACAGCAGCAGGCGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCACAGGTG	6189
QY	5983	AACAAGGAGAGGAGAGAGTCCCTCTCAAGTTTTCACGGTTAAACACTGTACCTATTAT	6042
Db	6190	AACAAGGCTGGGACCACTTTCAGTATAAATCTCAGCCAGAAAAACCAACTCTCTCATCT	6249
RESULT 6			
AX671105			
LOCUS			
Sequence 1 from Patent WO03004523.			
AX671105			
AX671105.1 GI:29329569			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1			
Zhu, Z.			
Regulation of human citron rho/rac-interacting kinase			
Patent: WO 03004523-A 1, 16-JAN-2003;			
Bayer Aktiengesellschaft (DE)			
LOCATION/Qualifiers			
1..6165			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
FEATURES			
source			
ORIGIN			
Query Match 92.0%; Score 5666.2; DB 6; Length 6165;			
Best Local Similarity 95.9%; Pred. No. 0;			
Matches 5947; Conservative 0; Mismatches 13; Indels 243; Gaps 3;			
QY	1	ATGTTGAAGTTCAAATATGGAGCGGGAATCCTTTGGATGCTGTGCTGCTGAACCCATT	60
Db	1	ATGTTGAAGTTCAAATATGGAGCGGGAATCCTTTGGATGCTGTGCTGCTGAACCCATT	60
QY	61	GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCAACCCCTTTATGACTCAA	120
Db	61	GCCAGCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCAACCCCTTTATGACTCAA	120
QY	121	CAGCAGATGTCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA	180
Db	121	CAGCAGATGTCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA	180
QY	181	GAAATCAGTCAAGCTGCTCTGATGAAGATTAAAGCAGTGAGCAACTTTGTCGGGAAGTAT	240
Db	181	GAAATCAGTCAAGCTGCTCTGATGAAGATTAAAGCAGTGAGCAACTTTGTCGGGAAGTAT	240
QY	241	TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAAAGGACTTCGAAGTACGA	300
Db	241	TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAAAGGACTTCGAAGTACGA	300
QY	301	AGTCTTTGAGTTGTGGTCACTTTCTGAGTGCAAGTGCTGTAAGAGAGAAAGCAACCGGG	360
Db	301	AGTCTTTGAGTTGTGGTCACTTTCTGAGTGCAAGTGCTGTAAGAGAGAAAGCAACCGGG	360
QY	361	GACATCTATGCTATCAAAAGTATGAAGAGAGGCTTTATTGGCCCGCAGGACAGTTTCA	420
Db	361	GACATCTATGCTATCAAAAGTATGAAGAGAGGCTTTATTGGCCCGCAGGACAGTTTCA	420
QY	421	TTTTTTGAGGAGAGCGGACATATATTCTCGAAGCACACAGCCGTGGATCCCCCAATTA	480
Db	421	TTTTTTGAGGAGAGCGGACATATATTCTCGAAGCACACAGCCGTGGATCCCCCAATTA	480

QY	481	CAGTATGCTTTTCCAGGACAAAATCACCTTTATCTGATGGAGGATATACAGCTGGAGGG	540
Db	481	CAGTATGCTTTTCCAGGACAAAATCACCTTTATCTGATGGAGGATATACAGCTGGAGGG	540
QY	541	GACTTGCTGTCACTTTTGAATAGATAGAGGACAGTATGATGATGATGATGATGATGATGAT	600
Db	541	GACTTGCTGTCACTTTTGAATAGATAGAGGACAGTATGATGATGATGATGATGATGATGAT	600
QY	601	TACTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTCATCTGATGGGATACGTGATCGA	660
Db	601	TACTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTCATCTGATGGGATACGTGATCGA	660
QY	661	GACATCAAGCTCAGAGACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT	720
Db	661	GACATCAAGCTCAGAGACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT	720
QY	721	GGATCTCCCGGAAAATGAATTCAAACAGATGGTGAATGCCAAACTCCCGATTTGGAGCC	780
Db	721	GGATCTCCCGGAAAATGAATTCAAACAGATGGTGAATGCCAAACTCCCGATTTGGAGCC	780
QY	781	CCAGATTACATGGCTCCTGAAAGTCTGACTGTGATGAACGGGATGGAAGGACACCTAC	840
Db	781	CCAGATTACATGGCTCCTGAAAGTCTGACTGTGATGAACGGGATGGAAGGACACCTAC	840
QY	841	GGCTGGAGCTGTGACTGGTGGTCACTGGGCGTGAATTCCTTATGAGATGATTTATGGAGA	900
Db	841	GGCTGGAGCTGTGACTGGTGGTCACTGGGCGTGAATTCCTTATGAGATGATTTATGGAGA	900
QY	901	TCGCCCTTCGACAGAGGAACTCTGCCAGAACCTTCAAATACATTTATGATTTCCAGCGG	960
Db	901	TCGCCCTTCGACAGAGGAACTCTGCCAGAACCTTCAAATACATTTATGATTTCCAGCGG	960
QY	961	TTTTTTGAAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1020
Db	961	TTTTTTGAAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1020
QY	1021	TTGTTGTGGGCGCAGAAAAGAGAGACTGAAGTTTGAAGGTTCTTTGCTGCCATCTCTTTCTC	1080
Db	1021	TTGTTGTGGGCGCAGAAAAGAGAGACTGAAGTTTGAAGGTTCTTTGCTGCCATCTCTTTCTC	1080
QY	1081	TCTAAATTTGACTGGAAACACATTCGTAACCTCTCCCTCCCTTCCTCCACCCCTCAAG	1140
Db	1081	TCTAAATTTGACTGGAAACACATTCGTAACCTCTCTCCCTCCCTTCCTCCACCCCTCAAG	1140
QY	1141	TCTGACGATGACACCTTCCAAATTTTGAAGAACACAGAGAGAAATTCGTGGGTTTCATCTCT	1200
Db	1141	TCTGACGATGACACCTTCCAAATTTTGAAGAACACAGAGAGAAATTCGTGGGTTTCATCTCT	1200
QY	1201	CCGTGCCAGTGAAGCTTCTCGGTTCTCGGTTGAAGAACTCCCTTTGTGGGTTTTCG	1260
Db	1201	CCGTGCCAGTGAAGCTTCTCGGTTCTCGGTTGAAGAACTCCCTTTGTGGGTTTTCG	1260
QY	1261	TACAGCAAGGCACTGGGGATTCTTTGGTGAATCTGAGTCTGTTGTGTCGGGTCTGGACTCC	1320
Db	1261	TACAGCAAGGCACTGGGGATTCTTTGGTGAATCTGAGTCTGTTGTGTCGGGTCTGGACTCC	1320
QY	1321	CTTGCCAAAGTACTAGCTCCATGGAAAAGAACTTCTCATCAAAAGCAAGAGCTTACAAGAC	1380
Db	1321	CTTGCCAAAGTACTAGCTCCATGGAAAAGAACTTCTCATCAAAAGCAAGAGCTTACAAGAC	1380
QY	1381	TCTCAGGACAAAGTGTCAAAAGATGAGGAGGAAATCACCGGTTTACATCGAGAGTGTCA	1440
Db	1381	TCTCAGGACAAAGTGTCAAAAGATGAGGAGGAAATCACCGGTTTACATCGAGAGTGTCA	1440
QY	1441	GAGGTGGAGGCTGTGTTAGTACAGAGGAGGTGGAGCTGAAAGCCCTCTGAGACTCAGAGA	1500
Db	1441	GAGGTGGAGGCTGTGTTAGTACAGAGGAGGTGGAGCTGAAAGCCCTCTGAGACTCAGAGA	1500
QY	1501	TCCTCTCTGGAGAGGAGGAGCTTCTACCTACATCAGAGATGAGTAAAGCGAAGT	1560
Db	1501	TCCTCTCTGGAGAGGAGGAGCTTCTACCTACATCAGAGATGAGTAAAGCGAAGT	1560
QY	1561	TTGGAGCAAGCAGGATGGAGGTGTCCCGAGGAGTGAACAAAGCACTGCAGCTTCTCCAT	1620



QY 3733 CAAACCAAACTCATTTGATTTTCTGCAAGCAAAATGGACCACCTGCTTAAAGAAAAAG 3792  
Db 3781 CAAACCAAACTCATTTGATTTTCTGCAAGCAAAATGGACCACCTGCTTAAAGAAAAA - - 3838  
QY 3793 GGTTTATTTAGTCGACGGAAGAGGACCTGCTTTACCCACACACAGGTTCTCTGCGATAC 3852  
Db 3839 -----AGGTTCTCTGCGATAC 3855  
QY 3853 AATGAGCTGAAGCTGGCCCTGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGAAAGCC 3912  
Db 3856 AATGAGCTGAAGCTGGCCCTGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGAAAGCC 3915  
QY 3913 CTTCAAGAGACCCGCATCGAGTCCGGTCCGCCCGGAGGAAGCTGCCCAACGCAAAACA 3972  
Db 3916 CTTCAAGAGACCCGCATCGAGTCCGGTCCGCCCGGAGGAAGCTGCCCAACGCAAAACA 3975  
QY 3973 AGGACCCACACCCATCCACGCGCAGCCAGCCGCGAGGACGACATGCCATGTCGCC 4032  
Db 3976 AGGACCCACACCCATCCACGCGCAGCCAGCCGCGAGGACGACATGCCATGTCGCC 4035  
QY 4033 ATCTGCGGTGCCAGAGACACAGCCAGTGCATGAGCTGCTGGCCCGGCATCCAGC 4092  
Db 4036 ATCTGCGGTGCCAGAGACACAGCCAGTGCATGAGCTGCTGGCCCGGCATCCAGC 4095  
QY 4093 CGCAGAAAGAGTCTTCAACTCCAGAGGAAATTTAGTGGCGTCTTAAAGAACGATGCAC 4152  
Db 4096 CGCAGAAAGAGTCTTCAACTCCAGAGGAAATTTAGTGGCGTCTTAAAGAACGATGCAC 4155  
QY 4153 CACAATATTCCTCACCGATTCAACGTAGGACTGAACATGCCAGCCACAAAGTGTCTGTG 4212  
Db 4156 CACAATATTCCTCACCGATTCAACGTAGGACTGAACATGCCAGCCACAAAGTGTCTGTG 4215  
QY 4213 TGTCTGGATACCGTGCATTTGGAGCGCAGGAGCATCCAAATGTCTCGAATGTCAGGTGATG 4272  
Db 4216 TGTCTGGATACCGTGCATTTGGAGCGCAGGAGCATCCAAATGTCTCGAATGTCAGGTGATG 4275  
QY 4273 TGTACCCCAAGTGTCCACGTGCTTGCAGACCACTGGCGTCTGCTGTAATATGCC 4332  
Db 4276 TGTACCCCAAGTGTCCACGTGCTTGCAGCCCACTGGCGTCTGCTGTAATATGCC 4335  
QY 4333 ACACACTTCACCGAGGCTTCTGCCGTGACAAATGAATCCCCAGGTTCCAGACCAG 4392  
Db 4336 ACACACTTCACCGAGGCTTCTGCCGTGACAAATGAATCCCCAGGTTCCAGACCAG 4395  
QY 4393 GAGCCAGCAGCAGCTTGACCTGGAGAGGTGATGAAGGTGCCAGGATTAACAAACA 4452  
Db 4396 GAGCCAGCAGCAGCTTGACCTGGAGAGGTGATGAAGGTGCCAGGATTAACAAACA 4455  
QY 4453 GGACAGCAAGGTGGACAGGAAGTACATTGTCTGGAGGATCAAAAGTCTCTATTAT 4512  
Db 4456 GGACAGCAAGGTGGACAGGAAGTACATTGTCTGGAGGATCAAAAGTCTCTATTAT 4515  
QY 4513 GACATGAAGCCAGAGAACTGGACAGAGCCGGTGGAAAGATTTGAGTGTGCTTCCC 4572  
Db 4516 GACATGAAGCCAGAGAACTGGACAGAGCCGGTGGAAAGATTTGAGTGTGCTTCCC 4575  
QY 4573 GAGCGGATGTATCTATTATCATGTGCGGTGCTTCCGAACTCGCAATACAGCCAAA 4632  
Db 4576 GAGCGGATGTATCTATTATCATGTGCGGTGCTTCCGAACTCGCAATACAGCCAAA 4635  
QY 4633 GCA----- 4635  
Db 4636 GCAGATGTCCATACATACTGAAGATGGAATCTCACCCGCAACACCTGCTGGCCCGG 4695  
QY 4636 ----- 4635  
Db 4696 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTACACCGCTTA 4755  
QY 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4662  
Db 4756 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAAACTG 4815

QY 4663 CTTGAAAACTCCCTGTGTGAAACTGGAAAGGTGATGACCGTCTAGACATGAATGACGCGCTG 4722  
Db 4816 CTTGAAAACTCCCTGTGTGAAACTGGAAAGGTGATGACCGTCTAGACATGAATGACGCGCTG 4875  
QY 4723 CCTTTCAGTCACAGGTTGTTGGTGGGCACCGAGGAAGGCTCTACGCCCTGGAATGTC 4782  
Db 4876 CCTTTCAGTCACAGGTTGTTGGTGGGCACCGAGGAAGGCTCTACGCCCTGGAATGTC 4935  
QY 4783 TTGAAAACTCCCTTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4842  
Db 4936 TTGAAAACTCCCTTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4995  
QY 4843 AAGACCTTGAGAAAGCTTACTCATGTAGCAGGAGAGAGCGGACACTGTCTCTGTGGAC 4902  
Db 4996 AAGACCTTGAGAAAGCTTACTCATGTAGCAGGAGAGAGCGGACACTGTCTCTGTGGAC 5055  
QY 4903 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCAGCCGACATCTCA 4962  
Db 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 5115  
QY 4963 CCCAACTTTTGAAGCTCAAGGGTGCACCTTGTGGGCGAGCAAGATGAGAAC 5022  
Db 5116 CCCAACTTTTGAAGCTCAAGGGTGCACCTTGTGGGCGAGCAAGATGAGAAC 5175  
QY 5023 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTGTCTTCTCCGCTTACAAGAAAAC 5082  
Db 5176 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTGTCTTCTCCGCTTACAAGAAAAC 5235  
QY 5083 CTCAGCAAAATATGCACTCCGAAAGAGATAGAGACCTCAGAGCCCTGCATCTCAAC 5142  
Db 5236 CTCAGCAAAATATGCACTCCGAAAGAGATAGAGACCTCAGAGCCCTGCATCTCAAC 5295  
QY 5143 TTCAACCAATTACAGTATCTCATTTGGAAACCAATAAATTTCTACGAAATCGACATGAACAG 5202  
Db 5296 TTCAACCAATTACAGTATCTCATTTGGAAACCAATAAATTTCTACGAAATCGACATGAACAG 5355  
QY 5203 TACACGCTCAGGAAATTCCTGGATAAGATGACCAATTCCTTGGCACCTGTGTGTTGCC 5262  
Db 5356 TACACGCTCAGGAAATTCCTGGATAAGATGACCAATTCCTTGGCACCTGTGTGTTGCC 5415  
QY 5263 GCCTTCTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAACAGCGAGGCGACGAGAG 5322  
Db 5416 GCCTTCTCCAAACAGCTTCCCTGTCTCAATCGTGAGTGAACAGCGAGGCGACGAGAG 5475  
QY 5323 GAGTACTGTGTGTTTCCAGMAATTTGGAGTGTCTGTGATTTCTTACGGAAGACGTAGC 5382  
Db 5476 GAGTACTGTGTGTTTCCAGMAATTTGGAGTGTCTGTGATTTCTTACGGAAGACGTAGC 5535  
QY 5383 CGCACAGACATCTCAAGTGGAGTCTTACCTTTGGCCTTTGCTTACAGAGAACCTTAT 5442  
Db 5536 CGCACAGACATCTCAAGTGGAGTCTTACCTTTGGCCTTTGCTTACAGAGAACCTTAT 5595  
QY 5443 CTGTTTGTGACCCACTTCACTCACTCGAAGTAAATGAGATCCAGGCACGCTCCTCAGCA 5502  
Db 5596 CTGTTTGTGACCCACTTCACTCACTCGAAGTAAATGAGATCCAGGCACGCTCCTCAGCA 5655  
QY 5503 GGGACCCCTGCCCAGAGGTACTTGGACATCCGAAACCCGGCTACCTGGGCCCTTGCCTAT 5562  
Db 5656 GGGACCCCTGCCCAGAGGTACTTGGACATCCGAAACCCGGCTACCTGGGCCCTTGCCTAT 5715  
QY 5563 TCCTCAGAGGGATTTACTTGGGTCTCTATACAGGATAAATTAAGGTCTATTGTGTC 5622  
Db 5716 TCCTCAGAGGGATTTACTTGGGTCTCTATACAGGATAAATTAAGGTCTATTGTGTC 5775  
QY 5623 AAGGGAACCTCTGTTGAAGAGTCCGGCACTGAACACCAACCCGGGCGCTCCACCTCCCGC 5682  
Db 5776 AAGGGAACCTCTGTTGAAGAGTCCGGCACTGAACACCAACCCGGGCGCTCCACCTCCCGC 5835  
QY 5683 AGCAGCCCAACAAAGCAGGAGCCCAACCGTACAAAGAGACATCATCAAGAGCGGTGCC 5742  
Db 5836 AGCAGCCCAACAAAGCAGGAGCCCAACCGTACAAAGAGACATCATCAAGAGCGGTGCC 5895  
QY 5743 TCCAGCCAGCGCCCGGAGGAGCCCGAGCCACCCGCGAGAGCCCAAGCACACCCACCGC 5802

```

|||||
5896 TCAGCCAGCGCGCCGAGGCGCCAGCCACCCGAGAGCAACCCACCCG 5955
QY TACCGAGGCGGACCGAGCTGCGAGGACAAAGTCTCTGGCGCCGCTGGAGCGA 5862
Db TACCGAGGCGGACCGAGCTGCGAGGACAAAGTCTCTGGCGCCGCTGGAGCGA 6015
QY GAGAAGTCCCGGCGGATGCTCAGCACGCGGAGAGCGGTCCCGGGAGGCTGTTT 5922
Db GAGAAGTCCCGGCGGATGCTCAGCACGCGGAGAGCGGTCCCGGGAGGCTGTTT 6075
QY GAAGACAGCAGCGGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCCAGGTG 5982
Db GAAGACAGCAGCGGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCCAGGTG 6135
QY AACAAAGGGAAGAGGCGAGAGTGC 6005
Db AACAAAGTCTGGGACGAGTCTTC 6158

RESULT 7
AX574425
LOCUS AX574425 6165 bp DNA linear PAT 07-JAN-2003
DEFINITION Sequence 1 from Patent WO02059325.
ACCESSION AX574425
VERSION AX574425.1 GI:27551751
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE 1
AUTHORS Yu, X., Miranda, M. and Friddle, C.J.
TITLE Human kinases and polynucleotides encoding the same
JOURNAL Patent: WO 02059325-A 1 01-AUG-2002;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..6165
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 91.9%; Score 5661.4; DB 6; Length 6165;
Best local Similarity 95.8%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

QY 1 ATGTTGAAGTTCAAATATGAGCGGGAATCCTTTGGATGCTGTGCTGAAACCCATT 60
Db 1 ATGTTGAAGTTCAAATATGAGCGGGAATCCTTTGGATGCTGTGCTGAAACCCATT 60
QY 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCCCTTTATGACTCAA 120
Db 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCCCTTTATGACTCAA 120
QY 121 CAGCAGATGTCCTCTTTCCGAGAAGGATATTAGATGCGCCTCTTTGTTCTTTGAA 180
Db 121 CAGCAGATGTCCTCTTTCCGAGAAGGATATTAGATGCGCCTCTTTGTTCTTTGAA 180
QY 181 GAATGAGTCAGCCTGCTCTGATGAGATTAGACGCTGAGCAACTTTGTCGGAGTAT 240
Db 181 GAATGAGTCAGCCTGCTCTGATGAGATTAGACGCTGAGCAACTTTGTCGGAGTAT 240
QY 241 TCCGACACCATAGCTAGTTACAGGAGTCCAGCCTTCGGCAAAGGACTTCGAACTCAGA 300
Db 241 TCCGACACCATAGCTAGTTACAGGAGTCCAGCCTTCGGCAAAGGACTTCGAACTCAGA 300
QY 301 AGCTTTGTTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 360
Db 301 AGCTTTGTTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 360
QY 361 GACATCTATGCTATGAAAGTATGAAAGAGAGGCTTTATTGGCCCGAGGAGGTTTCA 420

```

```

Db 361 GACATCTATGCTATGAAAGTATGAAAGAGAGGCTTTATTGGCCCGAGGAGCAGGTTTCA 420
QY 421 TTTTGTGAGGAGAGCGGAACATATTATCTCGAAGCAACAGCCGCTGGATCCCCCAATTA 480
Db 421 TTTTGTGAGGAGAGCGGAACATATTATCTCGAAGCAACAGCCGCTGGATCCCCCAATTA 480
QY 481 CAGTATGCTTTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540
Db 481 CAGTATGCTTTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540
QY 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACAGTATGATGAAACCTGATACAGTTT 600
Db 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACAGTATGATGAAACCTGATACAGTTT 600
QY 601 TACCTAGCTGAGTGAATTTTGGCTGTTTTCAGAGGCTTCATCTGATGGGATACGTCATCGA 660
Db 601 TACCTAGCTGAGTGAATTTTGGCTGTTTTCAGAGGCTTCATCTGATGGGATACGTCATCGA 660
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
Db 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720
QY 721 GGATCTGCGCGGAAATGAATTTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780
Db 721 GGATCTGCGCGGAAATGAATTTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780
QY 781 CCAGATTACATGCTCTGAAAGTCTGACTGTGATGAACGGGATGGAAGGACACCTAC 840
Db 781 CCAGATTACATGCTCTGAAAGTCTGACTGTGATGAACGGGATGGAAGGACACCTAC 840
QY 841 GGCTGCACTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 GGCTGCACTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 901 TCCCGCTTCGAGAGGAACTCTGCGAGAACTTCAATAACATTAATTAATTAATTAATTAATTA 960
Db 901 TCCCGCTTCGAGAGGAACTCTGCGAGAACTTCAATAACATTAATTAATTAATTAATTAATTA 960
QY 961 TTTTGAATTTCCAGATGACCCCAAGTCAGCAGTACCTTCTGATCTGATCTGATCTGATCTGAT 1020
Db 961 TTTTGAATTTCCAGATGACCCCAAGTCAGCAGTACCTTCTGATCTGATCTGATCTGATCTGAT 1020
QY 1021 TTGTTGTCGCGCCAGAAAGAGAGACTGAACTTTTGAAGTCTTTTGTCTGCTGCTGCTGCTGCT 1080
Db 1021 TTGTTGTCGCGCCAGAAAGAGAGACTGAACTTTTGAAGTCTTTTGTCTGCTGCTGCTGCTGCT 1080
QY 1081 TCTAAATTTGACTTGGAAACAACTTCCTCTCCCGCTTCGTTCCACCCCTCAAG 1140
Db 1081 TCTAAATTTGACTTGGAAACAACTTCCTCTCCCGCTTCGTTCCACCCCTCAAG 1140
QY 1141 TCTGACGATCACACCTTCCAATTTTGTATGAACACAGAGAATTCGTGGGTTTCATCCTCT 1200
Db 1141 TCTGACGATCACACCTTCCAATTTTGTATGAACACAGAGAATTCGTGGGTTTCATCCTCT 1200
QY 1201 CCGTGCCAGCTGAGCCCTCAGGCTCTCGGGTGAAGAACTGCGCTTTGTTGGGTTTTCG 1260
Db 1201 CCGTGCCAGCTGAGCCCTCAGGCTCTCGGGTGAAGAACTGCGCTTTGTTGGGTTTTCG 1260
QY 1261 TACAGCAAGGCTGAGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1320
Db 1261 TACAGCAAGGCTGAGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1320
QY 1321 CTTGCCAAGACTAGCTCCATGGAAGAAATCTTCTCATCAAAAGCAAGAGCTTACAAGAC 1380
Db 1321 CTTGCCAAGACTAGCTCCATGGAAGAAATCTTCTCATCAAAAGCAAGAGCTTACAAGAC 1380
QY 1381 TCTCAGCACAAGTGTCAAGATGAGGAGGAATCAACCGGTTATCGGAGAGTGTCA 1440
Db 1381 TCTCAGCACAAGTGTCAAGATGAGGAGGAATCAACCGGTTATCGGAGAGTGTCA 1440
QY 1441 GAGGTGAGGCTGTGCTTAGTACAGAGGAGGTGAGCTGAAGGCTCTGAGACTTCAGAGA 1500
Db 1441 GAGGTGAGGCTGTGCTTAGTACAGAGGAGGTGAGCTGAAGGCTCTGAGACTTCAGAGA 1500

```

QY 1501 TCCCTCTGGACGAGACCTTCTACCTACATCAGAAATGCAATGAGTAAAGCGAAGT 1560  
Db 1501 TCCCTCTGGACGAGACCTTCTACCTACATCAGAAATGCAATGAGTAAAGCGAAGT 1560  
QY 1561 TTGGAGCAAGCAGGATGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620  
Db 1561 TTGGAGCAAGCAGGATGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620  
QY 1621 GATATCAGAGACAGAGCCGAGCTCCAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680  
Db 1621 GATATCAGAGACAGAGCCGAGCTCCAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680  
QY 1681 GTGGAAAGAAATGAGTGTGATGATGATCAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAAGA 1740  
Db 1681 GTGGAAAGAAATGAGTGTGATGATGATCAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAAGA 1740  
QY 1741 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTGAAGAAATCAAG 1800  
Db 1741 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTGAAGAAATCAAG 1800  
QY 1801 CGGAAAGCCAGAGATGTCAGATCAATCTGTTGAAGGCTAAGGATCAAGGGAAGCTGAA 1860  
Db 1801 CGGAAAGCCAGAGATGTCAGATCAATCTGTTGAAGGCTAAGGATCAAGGGAAGCTGAA 1860  
QY 1861 GTGGAGAAATATCGGAAATCGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
Db 1861 GTGGAGAAATATCGGAAATCGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
QY 1921 CTCACAGAAATCGGAGAGGCT----- 1944  
Db 1921 CTCACAGAAATCGGAGAGGCT----- 1944  
QY 1945 -----GCAAGAGGAGCGCGGAGGAGCTGGAAGCCAGCGAGCCAGCGTGTGAG 1980  
Db 1945 -----GCAAGAGGAGCGCGGAGGAGCTGGAAGCCAGCGAGCCAGCGTGTGAG 1980  
QY 1981 AATATCCGCCAGGCAAGGAGCGGAGCGGAGGAGCTGGAAGCTGCAAGCCAGAG 2040  
Db 1981 AATATCCGCCAGGCAAGGAGCGGAGCGGAGGAGCTGGAAGCTGCAAGCCAGAG 2040  
QY 1993 GATTCCTTGAGGCAATCAGAAAGAGCTGTGGAAGCTGAGGAGCGCGGCTTCTCTG 2052  
Db 2041 GATTCCTTGAGGCAATCAGAAAGAGCTGTGGAAGCTGAGGAGCGCGGCTTCTCTG 2100  
QY 2053 GAGAACAAAGTAAAGAGATAGAGACCATGGAGCGTAGAGAAACAGACTGAAGGATGAC 2112  
Db 2101 GAGAACAAAGTAAAGAGATAGAGACCATGGAGCGTAGAGAAACAGACTGAAGGATGAC 2160  
QY 2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAATTTCTGGAGCTCGAAGAG 2172  
Db 2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAATTTCTGGAGCTCGAAGAG 2220  
QY 2173 AATATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG 2232  
Db 2221 AATATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG 2280  
QY 2233 CAGCACTATGAGGAAAGATTAAGTGTGGACAAATCAGATAAAGAAAGCTTGGCTGAC 2292  
Db 2281 CAGCACTATGAGGAAAGATTAAGTGTGGACAAATCAGATAAAGAAAGCTTGGCTGAC 2340  
QY 2293 AAGGAGACTGAGAAATGATGAGAGACAGAGGAGGAGCCCATGAGAGGCGCAA 2352  
Db 2341 AAGGAGACTGAGAAATGATGAGAGACAGAGGAGGAGCCCATGAGAGGCGCAA 2400  
QY 2353 ATTCTCAGGCAACAGAGGCGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2412  
Db 2401 ATTCTCAGGCAACAGAGGCGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2460  
QY 2413 CAGAGGATTTGGAATGCTGAGGCGAATPAACTTGCAGCAATAGCAGTCTTTTACC 2472  
Db 2461 CAGAGGATTTGGAATGCTGAGGCGAATPAACTTGCAGCAATAGCAGTCTTTTACC 2520  
QY 2473 CAAAGGAATGAGGCGCAGAGAGATGATTTCTGACTCAGGCAACAGAAATTTTAC 2532  
Db 2521 CAAAGGAATGAGGCGCAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580

QY 2533 CTGAGACACAGGCTGGAAAGTTGGAGGCCCAAGAACCGAAATCTGGAGGAGCAGCTGGAG 2592  
Db 2581 CTGAGACACAGGCTGGAAAGTTGGAGGCCCAAGAACCGAAATCTGGAGGAGCAGCTGGAG 2640  
QY 2593 AAGATCAGCCACCAAGCCACAGTGCAGAAATCGGCTGTGGAATCTGGAGACAAAGATTG 2652  
Db 2641 AAGATCAGCCACCAAGCCACAGTGCAGAAATCGGCTGTGGAATCTGGAGACAAAGATTG 2700  
QY 2653 CGGAGGTCAGTCTTAGAGCACGAGGAGCAGAAATCTGAGCTCAAGCCAGCTCAGCAGAG 2712  
Db 2701 CGGAGGTCAGTCTTAGAGCACGAGGAGCAGAAATCTGAGCTCAAGCCAGCTCAGCAGAG 2760  
QY 2713 CTACAGCTCTCCCTGAGGAGCGGAGTCACTGATGACAGCCCTGAGGCTGACCGGCG 2772  
Db 2761 CTACAGCTCTCCCTGAGGAGCGGAGTCACTGATGACAGCCCTGAGGCTGACCGGCG 2820  
QY 2773 GCCCTGAGAGCCAGCTTCGCCAGGCAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2832  
Db 2821 GCCCTGAGAGCCAGCTTCGCCAGGCAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2880  
QY 2833 GAAGAGGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGCCCAATTTGATGCT 2892  
Db 2881 GAAGAGGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGCCCAATTTGATGCT 2940  
QY 2893 CTTCTTAACAGCTGTACTGTAAATCAGAGCTGGAGGAGCAGCTAAACAGCTGACCGAG 2952  
Db 2941 CTTCTTAACAGCTGTACTGTAAATCAGAGCTGGAGGAGCAGCTAAACAGCTGACCGAG 3000  
QY 2953 GACAACCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3012  
Db 3001 GACAACCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060  
QY 3013 GGCCCAACAGCAGATTTGACAACTGCGAAGTGAAGTGACCACTCTCCGCCGAGATC 3072  
Db 3061 GGCCCAACAGCAGATTTGACAACTGCGAAGTGAAGTGACCACTCTCCGCCGAGATC 3120  
QY 3073 ACGGAAACAGAGATGTCAGCTTACCAGCCAGAAAGCAACAGCTGGAGGCTCTGAAGACCAAG 3132  
Db 3121 ACGGAAACAGAGATGTCAGCTTACCAGCCAGAAAGCAACAGCTGGAGGCTCTGAAGACCAAG 3180  
QY 3133 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGAGAA 3192  
Db 3181 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGAGAA 3240  
QY 3193 AAAAGGCGGAGCTGGAGGCGCTGGAGGAGCTCTCGGCTGATGAGAAATCCAGCTTTGAG 3252  
Db 3241 AAAAGGCGGAGCTGGAGGCGCTGGAGGAGCTCTCGGCTGATGAGAAATCCAGCTTTGAG 3300  
QY 3253 TGTCCGGTTCCAGAGCTGCAGAGAAATGCTGACACCGAGAAACAGAGCAGGCGGAGAGCC 3312  
Db 3301 TGTCCGGTTCCAGAGCTGCAGAGAAATGCTGACACCGAGAAACAGAGCAGGCGGAGAGCC 3360  
QY 3313 GATCAGCGGATCACCGAGTCTCGCCAGGTGAGCTGGAGTGAAGGAGCACAAGCT 3372  
Db 3361 GATCAGCGGATCACCGAGTCTCGCCAGGTGAGCTGGAGTGAAGGAGCACAAGCT 3420  
QY 3373 GAGATTCCTGCTCTGAGCAGGCTCTCAAGAGCAGAAAGCTGAAAGGCGGAGAGCTCTCT 3432  
Db 3421 GAGATTCCTGCTCTGAGCAGGCTCTCAAGAGCAGAAAGCTGAAAGGCGGAGAGCTCTCT 3480  
QY 3433 GACAAGCTCAATGACCTGGAGAAAGATGCTATGCTTGAATGAATGCCCGAAGCTTA 3492  
Db 3481 GACAAGCTCAATGACCTGGAGAAAGATGCTATGCTTGAATGAATGCCCGAAGCTTA 3540  
QY 3493 CAGCAGAGCTGGAGCTGACAGAGCTCAACAGAGGCTTCTGGAGAGCAGAGCCAAA 3552  
Db 3541 CAGCAGAGCTGGAGCTGACAGAGCTCAACAGAGGCTTCTGGAGAGCAGAGCCAAA 3600  
QY 3553 TTACAGCAGCAGATGAGCTGCAAGAAATCACAATTTCCGCTGACTCAAGAGCTGCAA 3612  
Db 3601 TTACAGCAGCAGATGAGCTGCAAGAAATCACAATTTCCGCTGACTCAAGAGCTGCAA 3660  
QY 3613 GAAGCTCTAGATCGGCTGATCTACTGAGACAGAAAGAGTACTGGAGTATCAGCTG 3672



QY 5683 AGCAGCCCCAACACGAGGCCCCACCCACGTAACAGAGGACATCACAAAGCGGTGGCC 5742  
Db 5836 ACCAGCCCCAACACGAGGCCCCACCCACGTAACAGAGGACATCACAAAGCGGTGGCC 5895  
QY 5743 TCCAGCCAGCGCCCGCGAGGCCCCAGCCACCGCGAGAGGCAAGACACACCCACCGC 5802  
Db 5896 TCCAGCCAGCGCCCGCGAGGCCCCAGCCACCGCGAGAGGCAAGACACACCCACCGC 5955  
QY 5803 TACCAGGAGGCGGAGCGAGCTGGCGAGGGAACAAGTCTCTGGCGGCCCTCTGAGGGA 5862  
Db 5956 TACCAGGAGGCGGAGCGAGCTGGCGAGGGAACAAGTCTCTGGCGGCCCTCTGAGGGA 6015  
QY 5863 GAGAGTCTCCCGCGCGAGTGTCTAGCAGCGGAGAGCGGTCCCGCGGAGCGTGT 5922  
Db 6016 GAGAGTCTCCCGCGCGAGTGTCTAGCAGCGGAGAGCGGTCCCGCGGAGCGTGT 6075  
QY 5923 GAAGCAGCAGCAGGCGCGGTGGCTGGCGGAGCGGTGAGGACCCCGCTGTCCAGGTG 5982  
Db 6076 GAAGCAGCAGCAGGCGCGGTGGCTGGCGGAGCGGTGAGGACCCCGCTGTCCAGGTG 6135  
QY 5983 AACAAAGGGAAGGCGAGAGTGC 6005  
Db 6136 AACAAAGTCTGGGACCAAGTCTTC 6158

RESULT 8  
AXI166510  
LOCUS AXI166510 6159 bp DNA linear PAT 22-JUN-2001  
DEFINITION Sequence 1 from Patent WO0138503.  
ACCESSION AXI166510  
VERSION AXI166510.1 GI:14546855  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Plowman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R.,  
Flanagan, P. and Clary, D.S.  
TITLE Novel human protein kinases and protein kinase-like enzymes  
JOURNAL Patent: WO 0138503-A 1 31-MAY-2001;  
Sugen, Inc. (US)  
FEATURES  
source  
1. 6159  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 91.7%; Score 5650.4; DB 6; Length 6159;  
Best Local Similarity 95.9%; Pred. No. 0;  
Matches 5943; Conservative 0; Mismatches 11; Indels 246; Gaps 4;  
QY 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGGTCTGCTGAACCCATT 60  
Db 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGGTCTGCTGAACCCATT 60  
QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGCTTCCAGGGGAACACCCCTTTATGACTCAA 120  
Db 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAACACCCCTTTATGACTCAA 120  
QY 121 CAGCAGATGTCTCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTCTCTTTGAA 180  
Db 121 CAGCAGATGTCTCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTCTCTTTGAA 180  
QY 181 GAATGAGTCAAGCTGTCTGATGAAGATTAAAGCAGTGAGCACTTTGTCCGGAA--G 237  
Db 181 GAATGAGTCAAGCTGTCTGATGAAGATTAAAGCAGTGAGCACTTTGTCCGGAAAGTG 240  
QY 238 TATTCCGACACCATAGTGTGAGTTACAGAGGCTCCAGGCTTCGGCAAAAGGACTTCGAAGTC 297  
Db 241 TATTCCGACACCATAGTGTGAGTTACAGAGGCTCCAGGCTTCGGCAAAAGGACTTCGAAGTC 300

QY 298 AGAAGTCTTGTAGTGTGTGCTGCTGAAAGTGCAGGTGCTAAAGAGAAAGCAACC 357  
Db 301 AGAAGTCTTGTAGTGTGTGCTGCTGAAAGTGCAGGTGCTAAAGAGAAAGCAACC 360  
QY 358 GGGGACATCTATGCTATGAAGTGAAGAGAAAGGCTTTATTGGCCCCAGGAGCAGTT 417  
Db 361 GGGGACATCTATGCTATGAAGTGAAGAGAAAGGCTTTATTGGCCCCAGGAGCAGTT 420  
QY 418 TCATTTTGGAGAAAGCGGAACATATTAATCTCGAAGCAACAGCCCGTGGATCCCCCAA 477  
Db 421 TCATTTTGGAGAAAGCGGAACATATTAATCTCGAAGCAACAGCCCGTGGATCCCCCAA 480  
QY 478 TTACAGTATCCCTTTTACAGGCAAAAATCACTTTATCTGATGGAGGAATATCAGCCTTGA 537  
Db 481 TTACAGTATCCCTTTTACAGGCAAAAATCACTTTATCTGATGGAGGAATATCAGCCTTGA 540  
QY 538 GGGGACTTGTGTCACTTTTGAATAGATATGAGGACCAAGTATGATAAAACCTTGATACAG 597  
Db 541 GGGGACTTGTGTCACTTTTGAATAGATATGAGGACCAAGTATGATAAAACCTTGATACAG 600  
QY 598 TTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCACTGATGGATAGCTGCAT 657  
Db 601 TTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCACTGATGGATAGCTGCAT 660  
QY 658 CGAGACATCAAGCCTGAGAAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGTGGAT 717  
Db 661 CGAGACATCAAGCCTGAGAAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGTGGAT 720  
QY 718 TTTGGATCTCGCGGAAAAATGAATTTCAACAAGATGCTGAATGCCAAACCTCCGATTTGGG 777  
Db 721 TTTGGATCTCGCGGAAAAATGAATTTCAACAAGATGCTGAATGCCAAACCTCCGATTTGGG 780  
QY 778 ACCCGAGATTACATGGCTCTCTGAAAGTCTGACTGTGATGAACGGGGATGAAAAAGCACC 837  
Db 781 ACCCGAGATTACATGGCTCTCTGAAAGTCTGACTGTGATGAACGGGGATGAAAAAGCACC 840  
QY 838 TACGGCTTGAAGTGTGACTGTGTGCTGAGTGGCGGTGATTCCTATGATGATGATTTATGGG 897  
Db 841 TACGGCTTGAAGTGTGACTGTGTGCTGAGTGGCGGTGATTCCTATGATGATGATTTATGGG 900  
QY 898 AGATCCCGCTTCGAGAGGGAACCTTCCAGAACCTTCAATTAACATTAATGAATTTCCAG 957  
Db 901 AGATCCCGCTTCGAGAGGGAACCTTCCAGAACCTTCAATTAACATTAATGAATTTCCAG 960  
QY 958 CGGTTTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATTCATCAA 1017  
Db 961 CGGTTTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATTCATCAA 1020  
QY 1018 AGCTTGTGTGCGGCCGAGAAAGAGAGACTGAAGTTTGAAGTCTTCTGCTGCCATCCCTTC 1077  
Db 1021 AGCTTGTGTGCGGCCGAGAAAGAGAGACTGAAGTTTGAAGTCTTCTGCTGCCATCCCTTC 1080  
QY 1078 TTCTCTAAATTTGACTGGAAACAATTCGTAACCTCTCTCCCGCTTCGTTCCACCCCTC 1137  
Db 1081 TTCTCTAAATTTGACTGGAAACAATTCGTAACCTCTCTCCCGCTTCGTTCCACCCCTC 1140  
QY 1138 AAGTCTGACGATGACACTCCCAATTTTGAAGAACACAGAGAAAGATTCGTGGGTTTCATCC 1197  
Db 1141 AAGTCTGACGATGACACTCCCAATTTTGAAGAACACAGAGAAAGATTCGTGGGTTTCATCC 1200  
QY 1198 TCTCCGTGCCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCCGTTTGTGGGGTTT 1257  
Db 1201 TCTCCGTGCCAGCTGAGCCCTCAGGCTTCTCGGCTGAGAACTGCCGTTTGTGGGGTTT 1260  
QY 1258 TCGTACAGCAAGGCACTGGGGATTCTTGGTAGATCTGAGTCTGTGTGTGGTCTGGTAC 1317  
Db 1261 TCGTACAGCAAGGCACTGGGGATTCTTGGTAGATCTGAGTCTGTGTGTGGTCTGGTAC 1320  
QY 1318 TCCCTGCGCAAGACTAGCTCCATGGAAGAAATTTCTCATCAAAAGCAAGAGACTACAA 1377  
Db 1321 TCCCTGCGCAAGACTAGCTCCATGGAAGAAATTTCTCATCAAAAGCAAGAGACTACAA 1380

QY 1378 GACTCTCAGGACAAAGTGTCAACAAGATGGAGCAGGAAATGACCCGGTTACATCGAGAGTG 1437  
DB 1381 GACTCTCAGGACAAAGTGTCAACAAGATGGAGCAGGAAATGACCCGGTTACATCGAGAGTG 1440  
QY 1438 TCAGAGGTGGAGGTGTGCTTAGTTCAGAGAGAGGTGGAGTGAAGCCCTCTGAGACTCAG 1497  
DB 1441 TCAGAGGTGGAGGTGTGCTTAGTTCAGAGAGAGGTGGAGTGAAGCCCTCTGAGACTCAG 1500  
QY 1498 AGATCCCTCTGGAGCAGGACCTTGCTTACATACACAGAAATGACAGTACGAGTAAAGCGA 1557  
DB 1501 AGATCCCTCTGGAGCAGGACCTTGCTTACATACACAGAAATGACAGTACGAGTAAAGCGA 1560  
QY 1558 AGTTTGAGCAAGCACCGATGGAGGTGTCCAGAGAGGATGACAAAGCACTGCGAGCTTCTC 1617  
DB 1561 AGTTTGAGCAAGCACCGATGGAGGTGTCCAGAGAGGATGACAAAGCACTGCGAGCTTCTC 1620  
QY 1618 CATGATATCAGAGCAGAGCCGAGCTCCAGAGAAATCAAAGACAGAGTACCAAGGCT 1677  
DB 1621 CATGATATCAGAGCAGAGCCGAGCTCCAGAGAAATCAAAGACAGAGTACCAAGGCT 1680  
QY 1678 CAAGTGAAGAAATGAGGTTGATGATCAATCACTGTGAGAGGATCTTGCTCAGCAAGA 1737  
DB 1681 CAAGTGAAGAAATGAGGTTGATGATCAATCACTGTGAGAGGATCTTGCTCAGCAAGA 1740  
QY 1738 AGACGGAGTATCTTACGAAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAGAAATTC 1797  
DB 1741 AGACGGAGTATCTTACGAAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAGAAATTC 1800  
QY 1798 AAGCGGAAACGACAGAAATGTACAGTAAACTCTTGAAGCTAAAGATCAAGGAGGCT 1857  
DB 1801 AAGCGGAAACGACAGAAATGTACAGTAAACTCTTGAAGCTAAAGATCAAGGAGGCT 1860  
QY 1858 GAGTGGAGAGATATGCGAAATCTGGAAGAGATCAATGCTGAGCAGAGCTCAAAATTCAG 1917  
DB 1861 GAGTGGAGAGATATGCGAAATCTGGAAGAGATCAATGCTGAGCAGAGCTCAAAATTCAG 1920  
QY 1918 GAGCTCCAGAGAAACTGGAGAGGCT----- 1944  
DB 1921 GAGCTCCAGAGAACTGGAGAGGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGTG 1980  
QY 1945 -----GCAAGAGCAGGAGCCGAGAGGAGCTGGAGAGCTGGAGAACTGCAGAACCGA 1989  
DB 1981 CAGAATATCCGACGCAAGAGAGCAGCGAGGAGCTGGAGAGCTGCAGAACCGA 2040  
QY 1990 GAGGATCTTCTGAAGGCATCAGAAAGAGCTGGTGAAGCTGAGGAAGCCGCCATTC 2049  
DB 2041 GAGGATCTTCTGAAGGCATCAGAAAGAGCTGGTGAAGCTGAGGAAGCCGCCATTC 2100  
QY 2050 CTGGAGAAACAAGTAAAGAGACTTAGAGACCATGGAGCGTAGAGAAAACAAGCTGAAGAT 2109  
DB 2101 CTGGAGAAACAAGTAAAGAGACTTAGAGACCATGGAGCGTAGAGAAAACAAGCTGAAGAT 2160  
QY 2110 GACATCCAGACAAATCCCAACAGATCCAGAGAGTGGCTGATAAAATCTGGAGCTCGAA 2169  
DB 2161 GACATCCAGACAAATCCCAACAGATCCAGAGAGTGGCTGATAAAATCTGGAGCTCGAA 2220  
QY 2170 GAGAAACATCGGAGGCCCAAGTCTCAGCCAGACCTAGAGTGCCTGAAAACAGAAA 2229  
DB 2221 GAGAAACATCGGAGGCCCAAGTCTCAGCCAGACCTAGAGTGCCTGAAAACAGAAA 2280  
QY 2230 GAGCAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCT 2289  
DB 2281 GAGCAGCACTATGAGGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGACCTGGCT 2340  
QY 2290 GACAGGAGACACTGAGAACATGATGCAGACACAGGAGGAGGCCATGAGAGGCG 2349  
DB 2341 GACAGGAGACACTGAGAACATGATGCAGACACAGGAGGAGGCCATGAGAGGCG 2400  
QY 2350 AAAATTTCTCAGGCAACAGAGGCGGATGATCAATGCTATGGATTCAGAGATCAGATCCCTG 2409  
DB 2401 AAAATTTCTCAGGCAACAGAGGCGGATGATCAATGCTATGGATTCAGAGATCAGATCCCTG 2460  
QY 2410 GACAGAGGATGTGGAAGTGTCTGAGGCCAATAAAGTGTGAGCAATATGAGCTCTTTT 2469

DB 2461 GAAACAGAGATTGTGAAACTGTCTGAAAGCCAAATAAATTGAGCAAAATAGCAGTCTTTT 2520  
QY 2470 ACCCAAAAGAAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTT 2529  
DB 2521 ACCCAAAAGAAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTT 2580  
QY 2530 TACCTGGAGACACAGGCTGGAGCTTGGAGCCCAAGAACCCGAAACTGGAGAGCAGCTG 2589  
DB 2581 TACCTGGAGACACAGGCTGGAGCTTGGAGCCCAAGAACCCGAAACTGGAGAGCAGCTG 2640  
QY 2590 GAGAAATCAGCCACCAAGACACAGTGCAGAGATCGGCTGCTGGAATCGGAACTGGAGACAAGA 2649  
DB 2641 GAGAAATCAGCCACCAAGACACAGTGCAGAGATCGGCTGCTGGAATCGGAACTGGAGACAAGA 2700  
QY 2650 TTGCGGAGGTCAGTCTTAGAGACAGAGAGCAGAAAATGGAGCTCAAGCGCCAGCTCA 2709  
DB 2701 TTGCGGAGGTCAGTCTTAGAGACAGAGAGCAGAAAATGGAGCTCAAGCGCCAGCTCA 2760  
QY 2710 GAGCTACAGCTCTCCCTCAGAGCGCGAGTCAAGTTGACAGCCCTGAGGCTGCACGG 2769  
DB 2761 GAGCTACAGCTCTCCCTCAGAGCGCGAGTCAAGTTGACAGCCCTGAGGCTGCACGG 2820  
QY 2770 GCGGCCCTGGAGAGCAGCTTTCGCGAGCGAAGACAGAGCTGGAGAGACACACAGCAGAA 2829  
DB 2821 GCGGCCCTGGAGAGCAGCTTTCGCGAGCGAAGACAGAGCTGGAGAGACACACAGCAGAA 2880  
QY 2830 GCTGAAGAGGAGATCCAGGCACTCAGCGCATAGAGATGAAATCCAGCGCAATTTGAT 2889  
DB 2881 GCTGAAGAGGAGATCCAGGCACTCAGCGCATAGAGATGAAATCCAGCGCAATTTGAT 2940  
QY 2890 GCTCTTCGTAAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTTAAACAGCTGACC 2949  
DB 2941 GCTCTTCGTAAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTTAAACAGCTGACC 3000  
QY 2950 GAGACAAACGCTGAACCTCAACACCAAAACTTCTACTTGTCCAAAACAACCTCGATGAGGCT 3009  
DB 3001 GAGACAAACGCTGAACCTCAACACCAAAACTTCTACTTGTCCAAAACAACCTCGATGAGGCT 3060  
QY 3010 TCTGGCCCAACAGCAGAGATTGTACAACTCGAAGTGAAGTGAACCTCTCCGCCGGAG 3069  
DB 3061 TCTGGCCCAACAGCAGAGATTGTACAACTCGAAGTGAAGTGAACCTCTCCGCCGGAG 3120  
QY 3070 ATCAGGAAACAGAGATGAGCTTACAGCCAGAGAAACGATGGAGGCTCTGAAGACC 3129  
DB 3121 ATCAGGAAACAGAGATGAGCTTACAGCCAGAGAAACGATGGAGGCTCTGAAGACC 3180  
QY 3130 ACCTGCACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGCTA 3189  
DB 3181 ACCTGCACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGCTA 3240  
QY 3190 GAAAAAGAGCGGAGCTGGAGGCTCTGGAGGAGGCTCTGGAGGAGGCTCTGGAGGAGGCTCT 3249  
DB 3241 GAAAAAGAGCGGAGCTGGAGGCTCTGGAGGAGGCTCTGGAGGAGGCTCTGGAGGAGGCTCT 3300  
QY 3250 GAGTGTGGGTTTCAGAGCTGAGAGATGCTGACACACAGAGAAACAGAGCAGGCGGAGA 3309  
DB 3301 GAGTGTGGGTTTCAGAGCTGAGAGATGCTGACACACAGAGAAACAGAGCAGGCGGAGA 3360  
QY 3310 GCCGATCAGCGGATCACCGAGTCTCCCGAGGCTGGAGGAGGCTGGAGGAGGAGCAAG 3369  
DB 3361 GCCGATCAGCGGATCACCGAGTCTCCCGAGGCTGGAGGAGGCTGGAGGAGGAGCAAG 3420  
QY 3370 GCTGAGATTTCTCCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGGCTC 3429  
DB 3421 GCTGAGATTTCTCCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGGCTC 3480  
QY 3430 TCTGACAAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAATGCTCCGAGC 3489  
DB 3481 TCTGACAAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAATGCTCCGAGC 3540  
QY 3490 TTACAGCAGAGCTGAGACTGAACAGAGCTCAACAGAGGCTTCTGAGAGAGCAGGCT 3549

Db 3541 TTACAGCAGAGCTGGAGACTGAACGAGAGCTCAACACAGAGGCTTCTGGAAGAGCAAGCC 3600  
QY 3550 AATATTACAGCAGAGATGGAACCTGCGAGAAAAATCAATTTTCGGTCTGACTCAAGAGACTG 3609  
Db 3601 AATATTACAGCAGAGATGGAACCTGCGAGAAAAATCAATTTTCGGTCTGACTCAAGAGACTG 3660  
QY 3610 CAAGAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGAGTATCAG 3669  
Db 3661 CAAGAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAATGAGTATCAG 3720  
QY 3670 CTGGAAAAATTCAGGTTCTCTATTCTCATGAAAAAGGTGAAATGGAAGGCACTATTCT 3729  
Db 3721 CTGGAAAAATTCAGGTTCTCTATTCTCATGAAAAAGGTGAAATGGAAGGCACTATTCT 3780  
QY 3730 CAACAACCAAACTCATTCGATTTCTGCAAGCCAAATGGAACCTCTCAAAAGAAA 3789  
Db 3781 CAACAACCAAACTCATTCGATTTCTGCAAGCCAAATGGAACCTCTCAAAAGAAA 3840  
QY 3790 AAGGGTTTATTAGTCGAGAAAGAGGACCTGCTTTTACCACACAGAGTTCCTCTGCGAG 3849  
Db 3841 A-----AGGTTCTCTGCGAG 3855  
QY 3850 TACAAATGAGCTGAAGTGGCCCTGGAGAGAGAGAAAGCTCGCTGTGCGAGAGCTAGAGAA 3909  
Db 3856 TACAATGAGCTGAAGTGGCCCTGGAGAGAGAGAAAGCTCGCTGTGCGAGAGCTAGAGAA 3915  
QY 3910 GCCCTTCAGAAAGACCCGCACTCGAGTCCGGTCCGCCGGGAGAAAGCTGCCACCGCAAA 3969  
Db 3916 GCCCTTCAGAAAGACCCGCACTCGAGTCCGGTCCGCCGGGAGAGCTGCCACCGCAAA 3975  
QY 3970 GMAAGGACCAACCAACCCATCCAGCCAGAGCAACCGGAGGAGAGAGATGCCATGTCC 4029  
Db 3976 GCAACGGAGCAACCAACCCATCCAGCCAGAGCAACCGGAGGAGAGATGCCATGTCC 4035  
QY 4030 GCCATCGTCGGTCCGAGAGACCAAGCCAGTGCCATGAGCTGTGTCGCCCGGCATCC 4089  
Db 4036 GCCATCGTCGGTCCGAGAGACCAAGCCAGTGCCATGAGCTGTGTCGCCCGGCATCC 4095  
QY 4090 AGCCGAGAAAGGAGTCTTCAATCCAGAGGAATTTAGTCGGCGTCTTAAAGAAACGCATG 4149  
Db 4096 AGCCGAGAAAGGAGTCTTCAATCCAGAGGAATTTAGTCGGCGTCTTAAAGAAACGCATG 4155  
QY 4150 CACCACAAATTCCTCACCAGATCAACGTAGACTGAACATGGAGCCACAAAGTGTGCT 4209  
Db 4156 CACCACAAATTCCTCACCAGATTCACGTAGACTGAACATGGAGCCACAAAGTGTGCT 4215  
QY 4210 GTGTGTCTGGATACCGTGCACATTTGGACGCCAGGATCCAAATGTCTCGAATGTCAAGTG 4269  
Db 4216 GTGTGTCTGGATACCGTGCACATTTGGACGCCAGGATCCAAATGTCTCGAATGTCAAGTG 4275  
QY 4270 ATGTGTCAACCAAGTGTCTCAGTGTCTCAGCCACCTCGCGGTGTCCTGCTGAATAT 4329  
Db 4276 ATGTGTCAACCAAGTGTCTCAGTGTCTTCCAGCCACCTCGCGGTGTCCTGCTGAATAT 4335  
QY 4330 GCCACACACTTCAACGAGGCTTCTCCGTGACAAATGAATCCCCAGGCTCTCCAGACC 4389  
Db 4336 GCCACACACTTCAACGAGGCTTCTCCGTGACAAATGAATCCCCAGGCTCTCCAGACC 4395  
QY 4390 AAGGAGCCAGCAGAGCTTGCACTCGAAGGAGGATGAAGTCCCGCAGGAATTAACAAA 4449  
Db 4396 AAGGAGCCAGCAGAGCTTGCACTCGAAGGAGGATGAAGTCCCGCAGGAATTAACAAA 4455  
QY 4450 CGAGGACAGAGGCTGGACAGAGATGATGTCTCGAGGGATCAAAATGCTCAT 4509  
Db 4456 CGAGGACAGAGGCTGGACAGAGATGATGTCTCGAGGGATCAAAATGCTCAT 4515  
QY 4510 TATGCAATGAAGCCAGAGAGCTGGACAGAGGCGGTGGAAGAAATTTGAGTGTGCTT 4569  
Db 4516 TATGCAATGAAGCCAGAGAGCTGGACAGAGGCGGTGGAAGAAATTTGAGTGTGCTT 4575  
QY 4570 CCGACGGGAGTGTATCTATTATGTGTGCTGTGCTTCCGAACCTCCCAAAATCAGCC 4629  
Db 4576 CCGACGGGAGTGTATCTATTATGTGTGCTGTGCTTCCGAACCTCCCAAAATCAGCC 4635

QY 4630 AAAGCAGA----- 4637  
Db 4636 AAAGCAGATGTCCCATACATACTGAAGATGAATCTCACCGCACAACAACCTGTGGCCC 4695  
QY 4638 ----- 4637  
Db 4696 GGGAGAACCCCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAAACAGCGCTGGGTCAACGCC 4755  
QY 4638 -----AAAGCAGAAGCTGATGCTAAA 4659  
Db 4756 TTAGAATCAGTTGTGCGAGGTGGAGAGTTTCTAGGAAAAAAGCAGAGTGTGCTAAA 4815  
QY 4660 CTGCTTTGAAACTCCCTGCTGAAACTGAAAGGTGATCACCGTCTAGACATGAATGCAAG 4719  
Db 4816 CTGCTTTGAAACTCCCTGCTGAAACTGAAAGGTGATCACCGTCTAGACATGAATGCAAG 4875  
QY 4720 CTGCCCTTCAAGTACACAGGTGTTTGGTGGGCAACCGAGAAAGGCTCTACGCCCTGAAT 4779  
Db 4876 CTGCCCTTCAAGTACACAGGTGTTTGGTGGGCAACCGAGAAAGGCTCTACGCCCTGAAT 4935  
QY 4780 GTCTTTGAAACTCCCTAAACCATGTCCAGGAAATTCGAGCAGTCTTCCAAATTTATAT 4839  
Db 4936 GTCTTTGAAACTCCCTAAACCATGTCCAGGAAATTCGAGCAGTCTTCCAAATTTATAT 4995  
QY 4840 ATCAAGGACCTGGAGAAAGCTACTCATATAGCAGAGAGAGCGGGCACTGTGTCTGTG 4899  
Db 4996 ATCAAGGACCTGGAGAAAGCTACTCATATAGCAGAGAGAGCGGGCACTGTGTCTGTG 5055  
QY 4900 GACGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGAGCATC 4959  
Db 5056 GACGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGAGCATC 5115  
QY 4960 TCACCCCAACATTTTGAAGCTGTCAAGGGTGTCCACATTTGTTGGGCGAGGCAAGATTGAG 5019  
Db 5116 TCACCCCAACATTTTGAAGCTGTCAAGGGTGTCCACATTTGTTGGGCGAGGCAAGATTGAG 5175  
QY 5020 AACGGGTCTGTGATCTGTGAGCCATGCCAGAGAAAGTGTGTCATCTCCCGCTACAAAGAA 5079  
Db 5176 AACGGGTCTGTGATCTGTGAGCCATGCCAGAGAAAGTGTGTCATCTCCCGCTACAAAGAA 5235  
QY 5080 AACCTCAGCAAAATCTGATCCGGAAGAGATAGAGACCTCAGAGCCCTGCGAGCTGTATC 5139  
Db 5236 AACCTCAGCAAAATCTGATCCGGAAGAGATAGAGACCTCAGAGCCCTGCGAGCTGTATC 5295  
QY 5140 CACTTCAACCAATTACAGTATCTCATTTGGAACCAATTAATTTACGAAATTCGACATGAAG 5199  
Db 5296 CACTTCAACCAATTACAGTATCTCATTTGGAACCAATTAATTTACGAAATTCGACATGAAG 5355  
QY 5200 CAGTACAGCTCGAGAAATTCCTGGATAAGATGACCACTTCTTGGCACCTGCTGTGTT 5259  
Db 5356 CAGTACAGCTCGAGAAATTCCTGGATAAGATGACCACTTCTTGGCACCTGCTGTGTT 5415  
QY 5260 GCCGCTCTTCCAAACAGCTTCCCTGTCTCAATCTGCAAGTGAACAGCGAGGCGAGCGA 5319  
Db 5416 GCCGCTCTTCCAAACAGCTTCCCTGTCTCAATCTGCAAGTGAACAGCGAGGCGAGCGA 5475  
QY 5320 GAGGAGTACTTGTGTTTCCAGAAATTTGGAGTGTTCGTGGAATTTTACGGAAGAGCT 5379  
Db 5476 GAGGAGTACTTGTGTTTCCAGAAATTTGGAGTGTTCGTGGAATTTTACGGAAGAGCT 5535  
QY 5380 AGCCGCAACAGAGATCTCAAGTGGAGTCCGTTACCTTTGGCTTTGCTTACAGAGAACCC 5439  
Db 5536 AGCCGCAACAGAGATCTCAAGTGGAGTCCGTTACCTTTGGCTTTGCTTACAGAGAACCC 5595  
QY 5440 TATCTGTTGTGACCCACTTCAACTACTCGAAGTAATGAGATTCAGGCAACCGTCTCTCA 5499  
Db 5596 TATCTGTTGTGACCCACTTCAACTACTCGAAGTAATGAGATTCAGGCAACCGTCTCTCA 5655  
QY 5500 GCAGGACCCCTGCGCAGAGGCTACCTGGACATCCGBAACCGGCTACCTGGGCCCTGCC 5559  
Db 5656 GCAGGACCCCTGCGCAGAGGCTACCTGGACATCCGBAACCGGCTACCTGGGCCCTGCC 5715

QY 5560 ATTTCTCAGAGGGGATTTACTTGGGCTCTCTATACAGGATAAATTAAGGTCATTTCG 5619  
Db 5716 ATTTCTCAGAGGGGATTTACTTGGGCTCTCTATACAGGATAAATTAAGGTCATTTCG 5775  
QY 5620 TGAAGGGAAACCTCTGTAAGGAGTCCGGCACTGAACACACCGGGGCCGCTCCACCTCC 5679  
Db 5776 TGAAGGGAAACCTCTGTAAGGAGTCCGGCACTGAACACACCGGGGCCGCTCCACCTCC 5835  
QY 5680 CGTAGAGCCCCCAACAGAGGAGCCACCCAGTCAACAGAGCAATCAACAGAGCGCTG 5739  
Db 5836 CGTAGAGCCCCCAACAGAGGAGCCACCCAGTCAACAGAGCAATCAACAGAGCGCTG 5895  
QY 5740 GCTTCCAGCCCCAGCGCCCGGAGGCGCCAGACCCCGGAGAGCAACACCCAC 5799  
Db 5896 GCTTCCAGCCCCAGCGCCCGGAGGCGCCAGACCCCGGAGAGCAACACCCAC 5955  
QY 5800 CGTACCGAGGGGGGACCGAGCTCGCGAGGCAAAAGTCTCTGGCCGCCCTCGGAG 5859  
Db 5956 CGTACCGAGGGGGGACCGAGCTCGCGAGGCAAAAGTCTCTGGCCGCCCTCGGAG 6015  
QY 5860 CGAGAGAGTCCCGCCGCGGATGCTCAGCACCGGAGAGCGGTCGCCCGGGAGCTG 5919  
Db 6016 CGAGAGAGTCCCGCCGCGGATGCTCAGCACCGGAGAGCGGTCGCCCGGGAGCTG 6075  
QY 5920 TTTGAAGACAGCAGCGGCGCGCTGCCCTGCGGAGCGCTGAGGACCCCGCTGCCAG 5979  
Db 6076 TTTGAAGACAGCAGCGGCGCGCTGCCCTGCGGAGCGCTGAGGACCCCGCTGCCAG 6135  
QY 5980 GTGAACAGGAGAGGGCA 5999  
Db 6136 GTGAACAGGAGGAGCA 6155

RESULT 9  
AX503780 6189 bp DNA linear PAT 27-SEP-2002  
LOCUS  
DEFINITION Sequence 10 from Patent WO0226826.  
ACCESSION AX503780  
VERSION AX503780.1 GI:23385965  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Gerlach, V.L., Macdougall, J.R., Smithson, G., Millet, I., Stone, D.,  
Gunther, E., Ellerman, K., Grosse, W.M., Alsbrook, J.P., Lepley, D.M.,  
Burgess, C.E., Padigar, M., Kekuda, R., Spytek, K.A., Leach, M.D. and  
Shimkets, R.A.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0226826-A 10 04-APR-2002;  
Curagen Corporation (US)  
FEATURES  
source  
1. .6189  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 91.4%; Score 5631; DB 6; Length 6189;  
Best Local Similarity 95.3%; Pred. No. 0;  
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;  
QY 1 ATGTTGAAGTTCAAATATGAGGCGCGGAATCTTTGGATGCTGGTCTGCTGAACCCATT 60  
Db 1 ATGTTGAAGTTCAAATATGAGGCGCGGAATCTTTGGATGCTGGTCTGCTGAACCCATT 60  
QY 61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCCCTTTATGACTCAA 120  
Db 61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCCCTTTATGACTCAA 120  
QY 121 CAGCAGATGTCCTCTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180

Db 121 CAGCAGATGTCCTCTCTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
QY 181 GAATGCACTAGCCTGCTCTGATGAAGATTAAAGCACTGAGCAACTTTGTCGGAAGTAT 240  
Db 181 GAATGCACTAGCCTGCTCTGATGAAGATTAAAGCACTGAGCAACTTTGTCGGAAGTAT 240  
QY 241 TCCGACACCATAGCTGAGTTACGAGGCTCCAGCTTCGGCAAGAGACTTCGAAAGTCAGA 300  
Db 241 TCCGACACCATAGCTGAGTTACGAGGCTCCAGCTTCGGCAAGAGACTTCGAAAGTCAGA 300  
QY 301 AGTCTTGTAGGTTGTGTCACATTTTCTGCTGAAAGTCAGGTGTTAAGAGAGAAAGCAAC 360  
Db 301 AGTCTTGTAGGTTGTGTCACATTTTCTGCTGAAAGTCAGGTGTTAAGAGAGAAAGCAAC 360  
QY 361 GACATCTATGCTATGAAGTGAAGAGAGGCTTTATTTGGCCAGGAGAGGTTTCA 420  
Db 361 GACATCTATGCTATGAAGTGAAGAGAGGCTTTATTTGGCCAGGAGAGGTTTCA 420  
QY 421 TTTTTCGAGGAGAGCGGAACATATTATCTCGAAGCACAAGCCGCTGGATCCCCCAATTA 480  
Db 421 TTTTTCGAGGAGAGCGGAACATATTATCTCGAAGCACAAGCCGCTGGATCCCCCAATTA 480  
QY 481 CAGTATGCTTTTTCAGGACAAAATCACCTTTATCTGATGAGAGATATCAGCTGAGGG 540  
Db 481 CAGTATGCTTTTTCAGGACAAAATCACCTTTATCTGATGAGAGATATCAGCTGAGGG 540  
QY 541 GACTTGTCTGTCACCTTTTGAATAGATAGAGACCAAGTGTAGTGAACCTTCCGATCAGT 600  
Db 541 GACTTGTCTGTCACCTTTTGAATAGATAGAGACCAAGTGTAGTGAACCTTCCGATCAGT 600  
QY 601 TACCTAGCTAGCTGATTTTGGCTGTTTCA CAGGCTTCATCTGATGGATACGTGATCGA 660  
Db 601 TACCTAGCTAGCTGATTTTGGCTGTTTCA CAGGCTTCATCTGATGGATACGTGATCGA 660  
QY 661 GACATCAAGCTGAGACATTTCTGTTACCGCAGCAGGACACATCAAGCTGGTGGATTT 720  
Db 661 GACATCAAGCTGAGACATTTCTGTTACCGCAGCAGGACACATCAAGCTGGTGGATTT 720  
QY 721 GGATCTCGCGGAAATGAATCAACAAAGATGGTGAATCCCAAACTCCGATGGGACC 780  
Db 721 GGATCTCGCGGAAATGAATCAACAA -- -GGTGAATGCCAACTCCCGATTTGGACC 777  
QY 781 CCAGATTACATGCTCTGTAAGTCTGACTGTGATGTAACCGGGATGGAAGAGCACCTAC 840  
Db 778 CCAGATTACATGCTCTGTAAGTCTGACTGTGATGTAACCGGGATGGAAGAGCACCTAC 837  
QY 841 GGCCTGACCTGTGACTGGTGTGAGTGGGCTGATTCCTTATGAGATGATTTATGGGAGA 900  
Db 838 GGCCTGACCTGTGACTGGTGTGAGTGGGCTGATTCCTTATGAGATGATTTATGGGAGA 897  
QY 901 TCCCGCTTCGAGAGGAACTCTGCGCAGAACCTTCAATTAACATATGAATTTCCAGCGG 960  
Db 898 TCCCGCTTCGAGAGGAACTCTGCGCAGAACCTTCAATTAACATATGAATTTCCAGCGG 957  
QY 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGTGATCTTCTGATCTGATTTCAAGC 1020  
Db 958 TTTTGTGAATTTCCAGATGACCCCAAGTGAAGTGTGATCTTCTGATCTGATTTCAAGC 1017  
QY 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGTCGCTATCTTTCTTC 1080  
Db 1018 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGTCGCTATCTTTCTTC 1077  
QY 1081 TCTAAATTTGACTGGAACAACTTCTGTAACCTCTCCCGCCCTTCGTTCCACCCCTCAAG 1140  
Db 1078 TCTAAATTTGACTGGAACAACTTCTGTAACCTCTCTCCCGCCCTTCGTTCCACCCCTCAAG 1137  
QY 1141 TCTGACGATGACCTCCAAATTTTGTGATGAACAGAGAGAAATTCGTGGGTTTCATCTCT 1200  
Db 1138 TCTGACGATGACCTCCAAATTTTGTGATGAACAGAGAGAAATTCGTGGGTTTCATCTCT 1197  
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTCGGTTTGTGGGTTTTCG 1260  
Db 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTCGGTTTGTGGGTTTTCG 1257



Db 3418 GAGATTCTCGCTCTGCGAGCGCTCTCAAGAGCAGAACTGAGGCCGAGAGCCTCTCT 3477  
Qy 3433 GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTTGAAATGAATGCCGGAAGCTTA 3492  
Db 3478 GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTTGAAATGAATGCCGGAAGCTTA 3537  
Qy 3493 CACGAGAGCTGAGAGCTGAAAGAGAGCTCAACAGAGAGCTTCTGGAGAGCAGCCAAA 3552  
Db 3538 CACGAGAGCTGAGAGCTGAAAGAGAGCTCAACAGAGAGCTTCTGGAGAGCAGCCAAA 3597  
Qy 3553 TTACAGCAGCAGATGACCTCGAGAGAGAAATCAATTTCCGCTGCTGACTCAAGGACTGCAA 3612  
Db 3598 TTACAGCAGCAGATGACCTCGAGAGAGAAATCAATTTCCGCTGCTGACTCAAGGACTGCAA 3657  
Qy 3613 GAAGCTCTAGATCGGCTGATCTACTGAAGAGAGAAAGTGAATTTGGAGTATCAGCTG 3672  
Db 3658 GAAGCTCTAGATCGGCTGATCTACTGAAGAGAGAAAGTGAATTTGGAGTATCAGCTG 3717  
Qy 3673 GAAAGCAATTCAGGTTCTCTATTTCTATGAAAGAGTGAAGAGGCACTATTTCTCAA 3732  
Db 3718 GAAAGCAATTCAGGTTCTCTATTTCTATGAAAGAGTGAAGAGGCACTATTTCTCAA 3777  
Qy 3733 CAAGCAAACTCATTTCTGCAAGCAAAATGGAACCACTGCTTAAAGAGAGAAAG 3792  
Db 3778 CAAGCAAACTCATTTCTGCAAGCAAAATGGAACCACTGCTTAAAGAGAGAAAG -- 3835  
Qy 3793 GGTATTATTTAGTCGAGAGAGAGGACCTGCTTTACCCACACAGGTTCTCTGAGTAC 3852  
Db 3836 -----AGTGGCTCTGAGTAC 3852  
Qy 3853 AATGAGCTGAAGTGGCCCTGGAGAGAGAAAGCTCGCTGCGAGAGCTAGAGAGGCC 3912  
Db 3853 AATGAGCTGAAGTGGCCCTGGAGAGAGAAAGCTCGCTGCGAGAGCTAGAGAGGCC 3912  
Qy 3913 CTTGAGAGACCGGATCAGCTCGGTCGGCCGGGAGAGAGTCCGCCACCGCAAGCA 3972  
Db 3913 CTTGAGAGACCGGATCAGCTCGGTCGGCCGGGAGAGAGTCCGCCACCGCAAGCA 3972  
Qy 3973 ACGGACACCAACCCATCCACGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4032  
Db 3973 ACGGACACCAACCCATCCACGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4032  
Qy 4033 ATGTCGGTTCGCGAGAGCAACGAGCCAGTGCATGAGCTGCTGGCCCGCCGCTCCAGC 4092  
Db 4033 ATGTCGGTTCGCGAGAGCAACGAGCCAGTGCATGAGCTGCTGGCCCGCCGCTCCAGC 4092  
Qy 4093 CGCAGAAAGAGTCTCAATCCAGAGGAAATTTAGTCGGCGCTTAAAGGAGCAGTGCAC 4152  
Db 4093 CGCAGAAAGAGTCTCAATCCAGAGGAAATTTAGTCGGCGCTTAAAGGAGCAGTGCAC 4152  
Qy 4153 CACAATATTTCTCAGCGAATTCAGTGAAGTGAACATCGAGAGCAGCAAGTGTCTGTG 4212  
Db 4153 CACAATATTTCTCAGCGAATTCAGTGAAGTGAACATCGAGAGCAGCAAGTGTCTGTG 4212  
Qy 4213 TGTCTGGATACGTCGACCTTGGAGCCGAGGATCCAAATGCTCGAATGTGAGTGTGATG 4272  
Db 4213 TGTCTGGATACGTCGACCTTGGAGCCGAGGATCCAAATGCTCGAATGTGAGTGTGATG 4272  
Qy 4273 TGTCAACCCCAAGTCTCCAGCTGCTTGGCAGGACACCTGGGGTGTGCTGCTGAATATGCC 4332  
Db 4273 TGTCAACCCCAAGTCTCCAGCTGCTTGGCAGGACACCTGGGGTGTGCTGCTGAATATGCC 4332  
Qy 4333 ACACATTCACGAGGCTTCTGCGGTGACAAAATGAACCTCCGAGGCTCCAGAGCAAG 4392  
Db 4333 ACACATTCACGAGGCTTCTGCGGTGACAAAATGAACCTCCGAGGCTCCAGAGCAAG 4392  
Qy 4393 GAGCCAGAGCAGCTTGCACCTTGGAGGGTGGATGAAGTGCACGAGGAAATACAAAGCA 4452  
Db 4393 GAGCCAGAGCAGCTTGCACCTTGGAGGGTGGATGAAGTGCACGAGGAAATACAAAGCA 4452  
Qy 4453 GAGCAGCAAGGCTGAGAGAGGATGATGCTTCTGGAGGATCAAAAGTCTCTATTTAT 4512

Db 4453 GGACAGCAAGGCTGGGACAGGAAGTACATTGTCTCTGAGGGATCAAAAGTCTCTATTAT 4512  
Qy 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCGGTGAAGAAATTTGAGCTGTGCTTCCC 4572  
Db 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCGGTGAAGAAATTTGAGCTGTGCTTCCC 4572  
Qy 4573 GACGGGATGATCTATTTCAATGCTGCTGCTTCCGAACTCGCAATATACAGCAAA 4632  
Db 4573 GACGGGATGATCTATTTCAATGCTGCTGCTTCCGAACTCGCAATATACAGCAAA 4632  
Qy 4633 GCA----- 4635  
Db 4633 GCAGATGCCATACATATGAAGATGAATCTACCCGACACACCTGCTGGCCGGG 4692  
Qy 4636 ----- 4635  
Db 4693 AGAACCCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTACCGCTTA 4752  
Qy 4636 -----GAAAAACAGCAAGCTGATGCTAACTG 4662  
Db 4753 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGAAAAAAGCAGAACTGATGCTAACTG 4812  
Qy 4663 CTTGAAAACTCCCTGCTGAACTGGAAGTGAACCGTCTAGACATGAACTGACGCTG 4722  
Db 4813 CTTGAAAACTCCCTGCTGAACTGGAAGTGAACCGTCTAGACATGAACTGACGCTG 4872  
Qy 4723 CCCTTCAGTCAACAGGTGTTGTTGGTGGGCAACGAGAGGGCTCTACGCCCTGAAATGTC 4782  
Db 4873 CCCTTCAGTCAACAGGTGTTGTTGGTGGGCAACGAGAGGGCTCTACGCCCTGAAATGTC 4832  
Qy 4783 TTGAAAACTCCCTTAACCCATGTCACAGAAATTTGGAGCAGTCTTCAAATTTATATATC 4842  
Db 4933 TTGAAAACTCCCTTAACCCATGTCACAGAAATTTGGAGCAGTCTTCAAATTTATATATC 4992  
Qy 4843 AAGGACCTGAGAGAGTACTCATGATAGAGAGAGAGAGAGGGGCACTGTGTCTTGTGGAC 4902  
Db 4993 AAGGACCTGAGAGAGTACTCATGATAGAGAGTGAAGAGGGGCACTGTGTCTTGTGGAC 5052  
Qy 4903 GTGAAGAAATGAAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGGCCCGACATCTCA 4962  
Db 5053 GTGAAGAAATGAAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGGCCCGACATCTCA 5112  
Qy 4963 CCACAACTTTTGAAGCTGCAAGGCTGCACTGTTTGGGGCAGCAGAGATTGAGAAC 5022  
Db 5113 CCACAACTTTTGAAGCTGCAAGGCTGCACTGTTTGGGGCAGCAGAGATTGAGAAC 5172  
Qy 5023 GGGCTCTGCACTCTGTGAGGCAATGCCAGCAAAAGTGTCTATTTCCGCTACAAACGAAAC 5082  
Db 5173 GGGCTCTGCACTCTGTGAGGCAATGCCAGCAAAAGTGTCTATTTCCGCTACAAACGAAAC 5232  
Qy 5083 CTGAGCAAAATCTGCACTCCGGAAGAGATAGAGACCTCAGAGCCCTGACGCTGATCCAC 5142  
Db 5233 CTGAGCAAAATCTGCACTCCGGAAGAGATAGAGACCTCAGAGCCCTGACGCTGATCCAC 5292  
Qy 5143 TTACCAATTTACAGTATCTCTCATTTGGAACCAATAAATTTACGAAATTCGACATGAAGCAG 5202  
Db 5293 TTACCAATTTACAGTATCTCTCATTTGGAACCAATAAATTTACGAAATTCGACATGAAGCAG 5352  
Qy 5203 TACACGCTGAGGAAATCTCTGGATAGAAATGACATTTCTTTGGCACTGCTGTGTTGCC 5262  
Db 5353 TACACGCTGAGGAAATCTCTGGATAGAAATGACATTTCTTTGGCACTGCTGTGTTGCC 5412  
Qy 5263 GCCTCTTCCAAACAGTCTCCCTGCTCAATCGTGAAGTGAACGAGCGAGCGAGGAG 5322  
Db 5413 GCCTCTTCCAAACAGTCTCCCTGCTCAATCGTGAAGTGAACGAGCGAGCGAGGAG 5472  
Qy 5323 GAGTACTTGTGTTTCCAGAAATTTGGAGTGTGTTGGAATTTCTTACGGAAGACGTAGC 5382  
Db 5473 GAGTACTTGTGTTTCCAGAAATTTGGAGTGTGTTGGAATTTCTTACGGAAGACGTAGC 5532  
Qy 5383 CGCACAGACCATCTCAGTGGAGTCCGTTACCTTTGCGCTTTCCTACAGAGAACCTAT 5442  
Db 5533 CGCACAGACCATCTCAGTGGAGTCCGTTACCTTTGCGCTTTCCTACAGAGAACCTAT 5592

```
QY 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCAGGACAGCTCCTCAGCA 5502
Db 5593 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCAGGACAGCTCCTCAGCA 5652
QY 5503 GGGACCCCTGCGGAGCGTACCTGACATCCGAAACCCGCGTACTGGGCGCTGCCATT 5562
Db 5653 GGGACCCCTGCGGAGCGTACCTGACATCCGAAACCCGCGTACTGGGCGCTGCCATT 5712
QY 5563 TCCTCAGGAGCGATTACTTTGGCGTCTCATACAGGATAAAATTAAAGGTCATTTCCTGC 5622
Db 5713 TCCTCAGGAGCGATTACTTTGGCGTCTCATACAGGATAAAATTAAAGGTCATTTCCTGC 5772
QY 5623 AAGGNAACCTGTAAGAGTCCGCACTGAACACACCGGGCGGCTCCACCTCCCGC 5682
Db 5773 AAGGNAACCTGTAAGAGTCCGCACTGAACACACCGGGCGGCTCCACCTCCCGC 5832
QY 5683 AGCAGCCCAACAAAGGAGCGCCACCCAGTACAAGAGCACATCACCAAGCGCGTGCC 5742
Db 5833 AGCAGCCCAACAAAGGAGCGCCACCCAGTACAAGAGCACATCACCAAGCGCGTGCC 5892
QY 5743 TCAGCCAGCGCCCGCCGAGGCGCCAGCCACCGCGAGAGCCCAAGCACACCCACCGC 5802
Db 5893 TCAGCCAGCGCCCGCCGAGGCGCCAGCCACCGCGAGAGCCCAAGCACACCCACCGC 5952
QY 5803 TACCGGAGGCGGAGCGAGCTGGCAGGACAGTCTCTGGCGGCGCCCTGGAGCGA 5862
Db 5953 TACCGGAGGCGGAGCGAGCTGGCAGGACAGTCTCTGGCGGCGCCCTGGAGCGA 6012
QY 5863 GAGAAAGTCCCGCGCGAGTGTCTAGCAGCGCGAGAGCGGTCCCGCGGAGGCTGTT 5922
Db 6013 GAGAAAGTCCCGCGCGAGTGTCTAGCAGCGCGAGAGCGGTCCCGCGGAGGCTGTT 6072
QY 5923 GAAGCAGCAGAGGCGGCGCGTGTCTGGGAGCGCGTGAGGACCCCGTGTCCAGGTG 5982
Db 6073 GAAGCAGCAGAGGCGGCGCGTGTCTGGGAGCGCGTGAGGACCCCGTGTCCAGGTG 6132
QY 5983 AACAAAGGAGGAGGCGAGAGTGCCTCAAGTTTTCAGGTTTACAGTGTCACT 6037
Db 6133 AACAAAGTGTGGACCGAGTCTTCAGTATTAATCTCAGCCAGAAAACCACTCCT 6187

RESULT 10
AX503778
LOCUS AX503778 6201 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 8 from Patent WO0226826.
ACCESSION AX503778
VERSION AX503778.1 GI:23385964
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
MAMMALIA; EUTHERIA; PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
REFERENCE 1
AUTHORS Gerlach, V.L., Macdougall, J.R., Smithson, G., Millet, I., Stone, D.,
Gunther, E., Eilerman, K., Grosse, W.M., Alsobrook, J.P., Lepley, D.M.,
Burgess, C.B., Padigar, M., Keku, R., Spytek, K.A., Leach, M.D. and
Shinkets, R.A.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0226826-A 8 04-APR-2002;
FEATURES Curagen Corporation (US)
source Location/Qualifiers
1. .6201
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 91.4%; Score 5629.8; DB 6; Length 6201;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;
QY 1 ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCTGCTGTAACCCATT 60
```

```
Db 1 ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGGATGCTGCTGTAACCCATT 60
QY 61 GCCAGCGGGCCTCCAGGCTGAATCTCTTCCAGGGGAAACACACCTTTATGACTCAA 120
Db 61 GCCAGCGGGCCTCCAGGCTGAATCTCTTCCAGGGGAAACACACCTTTATGACTCAA 120
QY 121 CAGCAGATGCTCTCTTTCCGAGAGGGATATAGATGCCCTCTTTGTTCTCTTTGAA 180
Db 121 CAGCAGATGCTCTCTTTCCGAGAGGGATATAGATGCCCTCTTTGTTCTCTTTGAA 180
QY 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCACTTTGTCGGAAGTAT 240
Db 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCACTTTGTCGGAAGTAT 240
QY 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTTCCGCAAGGACTTTGAAAGTCAGA 300
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTTCCGCAAGGACTTTGAAAGTCAGA 300
QY 301 AGTCTTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGTGTGTAAGAGAAAGCAACCGG 360
Db 301 AGTCTTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGTGTGTAAGAGAAAGCAACCGG 360
QY 361 GACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTGGCCAGGAGCAGGTTTCA 420
Db 361 GACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTGGCCAGGAGCAGGTTTCA 420
QY 421 TTTTGTGAGGAAGCGGAAACATATATCTGAAAGCAAGAGCCCGTGGATCCCCCAATTA 480
Db 421 TTTTGTGAGGAAGCGGAAACATATATCTGAAAGCAAGAGCCCGTGGATCCCCCAATTA 480
QY 481 CAGTATGCTTTCAGGACAAATAATCAGCTTTATCTGATGAGGAATATCAGCTGGAGGG 540
Db 481 CAGTATGCTTTCAGGACAAATAATCAGCTTTATCTGATGAGGAATATCAGCTGGAGGG 540
QY 541 GACTTGTGTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600
Db 541 GACTTGTGTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600
QY 601 TACCTAGCTAGCTGATTTTGGCTGTTTCAAGGTTTCACTGATGAGTATGAGTGCATCGA 660
Db 601 TACCTAGCTAGCTGATTTTGGCTGTTTCAAGGTTTCACTGATGAGTATGAGTGCATCGA 660
QY 661 GACATCAAGCTTGAGAAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTT 720
Db 661 GACATCAAGCTTGAGAAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTT 720
QY 721 GGATCTGCGCGAAATGAATTCAAAAGATGGTGAATGCCAAACCTCCCGATTTGGGACC 780
Db 721 GGATCTGCGCGAAATGAATTCAAAAGATGGTGAATGCCAAACCTCCCGATTTGGGACC 777
QY 781 CCAGATTACATGCTCTGAGTGAAGTGTGATGAAACGGGATGAAAGGACACCTAC 840
Db 778 CCAGATTACATGCTCTGAGTGAAGTGTGATGAAACGGGATGAAAGGACACCTAC 837
QY 841 GGCTGGAAGTGTGACTGGTGTGATGGGCGTGAATGCTGATGAGTATGAGTATGGGAGA 900
Db 838 GGCTGGAAGTGTGACTGGTGTGATGGGCGTGAATGCTGATGAGTATGAGTATGGGAGA 897
QY 901 TCCCGCTTCGACAGGAGAACCTTCGACAGAACCTTCATTAACATTAATGAAATTTCCAGCG 960
Db 898 TCCCGCTTCGACAGGAGAACCTTCGACAGAACCTTCATTAACATTAATGAAATTTCCAGCG 957
QY 961 TTTTGTGAAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
Db 958 TTTTGTGAAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1017
QY 1021 TTGTTGTGCGGCAGAAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCTGCCATCCTTTCTTC 1080
Db 1018 TTGTTGTGCGGCAGAAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGCTGCCATCCTTTCTTC 1077
QY 1081 TCTAAATTTGACTGGAACAAACATTTGTAACCTCTCCCGCTTCGTTCCACCCCTCAAG 1140
```

Db 1078 TCTAAAATTGACTGGAACAACATTCGTAAACGCTCTCTCCCTCCCTCGTTCCCTCCCTCAAG 1137  
Qy 1141 TCTGACGATGACACCTCCCAATTTTGATGAACACAGAGAAATTCGTGGGTTTTCATCCCTCT 1200  
Db 1138 TCTGACGATGACACCTCCCAATTTTGATGAACACAGAGAAATTCGTGGGTTTTCATCCCTCT 1197  
Qy 1201 CCGTGCAGCTGAGCCCTCAGCTTCTCGGGTGAAGAACTGCGCGTTTGTGGGGTTTTCG 1260  
Db 1198 CCGTGCAGCTGAGCCCTCAGCTTCTCGGGTGAAGAACTGCGCGTTTGTGGGGTTTTCG 1257  
Qy 1261 TACAGCAAGGCACTGGGATCTTGTGTAGATCTGAGTCTGTTGTGTCGGGTCTGAGCTCC 1320  
Db 1258 TACAGCAAGGCACTGGGATCTTGTGTAGATCTGAGTCTGTTGTGTCGGGTCTGAGCTCC 1317  
Qy 1321 CCGTCCCAAGCTAGCTCCATCGGAAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380  
Db 1318 CCGTCCCAAGCTAGCTCCATCGGAAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1377  
Qy 1381 TCTCAGGACAAAGTGTCAAGATGGAGCGGAATGACCGGTTTACATCGGAGGTGCA 1440  
Db 1378 TCTCAGGACAAAGTGTCAAGATGGAGCGGAATGACCGGTTTACATCGGAGGTGCA 1437  
Qy 1441 GAGGTGGAGCTGTGTTAGTTCAGAAAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
Db 1438 GAGGTGGAGCTGTGTTAGTTCAGAAAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1497  
Qy 1501 TCCTCTCTGGAGCAGACCTTGCTACTATACACAGAAATGAGTAGCTTAAAGCGAAGT 1560  
Db 1498 TCCTCTCTGGAGCAGACCTTGCTACTATACACAGAAATGAGTAGCTTAAAGCGAAGT 1557  
Qy 1561 TTGAGCAAGCAGGATGAGGTGTCACAGGAGGATGACAAAGCACTGCGAGTTCTCCAT 1620  
Db 1558 TTGAGCAAGCAGGATGAGGTGTCACAGGAGGATGACAAAGCACTGCGAGTTCTCCAT 1617  
Qy 1621 GATPATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGACGAGGATCACAGGCTCAA 1680  
Db 1618 GATPATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGACGAGGATCACAGGCTCAA 1677  
Qy 1681 GTGGAAGAAATGAGTTGATGATGAATCAGTTGGAGAGGATCTGTCTCACAAGAGA 1740  
Db 1678 GTGGAAGAAATGAGTTGATGATGAATCAGTTGGAGAGGATCTGTCTCACAAGAGA 1737  
Qy 1741 CGGAGTGATCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1800  
Db 1738 CGGAGTGATCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATCAAG 1797  
Qy 1801 CGGAAGCAGCAATGTCAGCAATCACTGTTGAAGGCTAAGGATCAAGGAGGCTGAA 1860  
Db 1798 CGGAAGCAGCAATGTCAGCAATCACTGTTGAAGGCTAAGGATCAAGGAGGAGCTGAA 1857  
Qy 1861 GTGGGAGAAATATCGGAAAATGGAAGAGATCAATGCTGAGCAGAGCTCAAAAATCAGGAG 1920  
Db 1858 GTGGGAGAAATATCGGAAAATGGAAGAGATCAATGCTGAGCAGAGCTCAAAAATCAGGAG 1917  
Qy 1921 CTCACAGAGAACTGAGAGAGCT----- 1944  
Db 1918 CTCACAGAGAACTGAGAGAGCTGTAAAGGCGAGCAGCGAGGCGCCAGAGCTGCTGCAG 1977  
Qy 1945 -----GCAAGAGCGAGCGGAGGAGCTGGAAGAGCTGAGAACCGAGAG 1992  
Db 1978 AATATCGGCGAGCAAGGAGCGAGCGGAGGAGCTGAGAGAGCTGAGAACCGAGAG 2037  
Qy 1993 GATTTCTTGAAGGATACGAAGAGAGCTGGTGAAGAGCTGAGAAACGCGCCCATTTCTGT 2052  
Db 2038 GATTTCTTGAAGGATACGAAGAGAGCTGGTGAAGAGCTGAGAAACGCGCCCATTTCTGT 2097  
Qy 2053 GAGAACAGGTAAAGAGCTAGAGCCATGGAGCGGTAGAGAAACAGACTGAAGATGAC 2112  
Db 2098 GAGAACAGGTAAAGAGCTAGAGCCATGGAGCGGTAGAGAAACAGACTGAAGATGAC 2157  
Qy 2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCAAGAG 2172  
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCAAGAG 2217

Qy 2173 AAACATCGGAGGCCCAAGTCTCAGCCAGCACTAGAACTGCACTGTAAACAGAAAGAG 2232  
Db 2218 AAACATCGGAGGCCCAAGTCTCAGCCAGCACTAGAACTGCACTGTAAACAGAAAGAG 2277  
Qy 2233 CAGCACTATGAGAAAGATTAAGTCTGGAACAATCAGATTAAGAAAGACCTGGCTGAC 2292  
Db 2278 CAGCACTATGAGAAAGATTAAGTCTGGAACAATCAGATTAAGAAAGACCTGGCTGAC 2337  
Qy 2293 AAGGAGACACTGGAGAACATGATGACAGAGACAGAGGAGGAGGCCATGAGAAAGGCAAA 2352  
Db 2338 AAGGAGACACTGGAGAACATGATGACAGAGACAGAGGAGGAGGCCATGAGAAAGGCAAA 2397  
Qy 2353 ATTCTCAGCGAACAGAAAGGCGATGATCAATGCTATGATTCGAAGATCAGATCCCTGGAA 2412  
Db 2398 ATTCTCAGCGAACAGAAAGGCGATGATCAATGCTATGATTCGAAGATCAGATCCCTGGAA 2457  
Qy 2413 CAGAGGATTTGGAACCTGTCTGAAGCCAAATAAACTTGCAGCAATAGCAGTCTTTTACC 2472  
Db 2458 CAGAGGATTTGGAACCTGTCTGAAGCCAAATAAACTTGCAGCAATAGCAGTCTTTTACC 2517  
Qy 2473 CAAAGGAACTGAAAGGCCCAAGAGAGATGATTTTGAACCTCAGGCAACAGAAATTTTAC 2532  
Db 2518 CAAAGGAACTGAAAGGCCCAAGAGAGATGATTTTGAACCTCAGGCAACAGAAATTTTAC 2577  
Qy 2533 CTGGAGACACAGGCTGGAAAGTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTGGAG 2592  
Db 2578 CTGGAGACACAGGCTGGAAAGTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTGGAG 2637  
Qy 2593 AAGATCAGCCACCAAGACCAACAGTGAAGAACTCGGCTGCTGGAACCTGGAGACAAGATTG 2652  
Db 2638 AAGATCAGCCACCAAGACCAACAGTGAAGAACTCGGCTGCTGGAACCTGGAGACAAGATTG 2697  
Qy 2653 CGGAGGTCAGTCTAGACACAGGAGGAGCAGAACTGGAGCTCAAGCGCAGCTCAGAGAG 2712  
Db 2698 CGGAGGTCAGTCTAGACACAGGAGGAGCAGAACTGGAGCTCAAGCGCAGCTCAGAGAG 2757  
Qy 2713 CTACAGCTCTCCCTGACAGGAGCGAGTCAAGTTGACAGCCCTGACAGCTGACCGGCG 2772  
Db 2758 CTACAGCTCTCCCTGACAGGAGCGAGTCAAGTTGACAGCCCTGACAGCTGACCGGCG 2817  
Qy 2773 GCCTTGAGAGCCAGCTTCCGAGCGGAGACAGAGCTGGAAGAGACCAAGCAGAGCT 2832  
Db 2818 GCCTTGAGAGCCAGCTTCCGAGCGGAGACAGAGCTGGAAGAGACCAAGCAGAGCT 2877  
Qy 2833 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGGCAAAATTTGATGCT 2892  
Db 2878 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAAAATTTGATGCT 2937  
Qy 2893 CTTGCTAACAGCTGTACTGTAAATCAAGACTGGAGAGCAGCTAAACAGAGCTGACGAG 2952  
Db 2938 CTTGCTAACAGCTGTACTGTGATCAGAGCTGGAGAGCAGCTAAACAGAGCTGACGAG 2997  
Qy 2953 GACAACTGAACTCAACACCAAACTTCTACTGTGTCACAACTTCGATGAGGCTTCT 3012  
Db 2998 GACAACTGAACTCAACACCAAACTTCTACTGTGTCACAACTTCGATGAGGCTTCT 3057  
Qy 3013 GGCGCCCAACGAGATTTGTACAACTGCGAAGTGAAGTGGACCATCTCCCGCGGAGATC 3072  
Db 3058 GGCGCCCAACGAGATTTGTACAACTGCGAAGTGAAGTGGACCATCTCCCGCGGAGATC 3117  
Qy 3073 ACGGAAACAGAGATGAGCTTACCGCAGACCAACAGATGGAGCTCTGAGAGACAG 3132  
Db 3118 ACGGAAACAGAGATGAGCTTACCGCAGACCAACAGATGGAGCTCTGAGAGACAG 3177  
Qy 3133 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTTAAACAGATGAGCTGTAGAA 3192  
Db 3178 TGCAACCATGCTGGAGGAAACAGGTCATGATTTGGAGGCCCTTAAACAGATGAGCTGTAGAA 3237  
Qy 3193 AAGAGCGGAGTGGGAGGCTGAGAGAGCTCTGGGTGATGAGAAATCCGAGTTTGGAG 3252  
Db 3238 AAGAGCGGAGTGGGAGGCTGAGAGAGCTCTGGGTGATGAGAAATCCGAGTTTGGAG 3297

QY 3253 TGTGGGTTTCAGAGCTGCGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGAGAGCC 3312  
Db 3298 TGTGGGTTTCAGAGCTGCGAGAGTGTGAGAGTGTGAGACACCGAGAAACAGAGCAGGCGAGAGCC 3357  
QY 3313 GATCAGCGGATCACCGAGTCTCGCCAGGTGTGAGCTGGCAGTGAAGAGCACAAGGCT 3372  
Db 3358 GATCAGCGGATCACCGAGTCTCGCCAGGTGTGAGCTGGCAGTGAAGAGCACAAGGCT 3417  
QY 3373 GAGATTCTCGCTCTCGCAGCAGGCTCTCAAAGAGCAGAAAGCTGAAGCGCCGAGAGCTCTCT 3432  
Db 3418 GAGATTCTCGCTCTCGCAGCAGGCTCTCAAAGAGCAGAAAGCTGAAGCGCCGAGAGCTCTCT 3477  
QY 3433 GACAGCTCAATGACCTGGAGAGAGCAATGCTATGCTTGAATGAATGATCCCGAGCTTA 3492  
Db 3478 GACAGCTCAATGACCTGGAGAGAGCAATGCTATGCTTGAATGAATGATCCCGAGCTTA 3537  
QY 3493 CAGCAGAGCTGGAGCTGAACAGAGCTCAAACAGAGGCTTCTGGAAGAGCAGACCCAAA 3552  
Db 3538 CAGCAGAGCTGGAGCTGAACAGAGCTCAAACAGAGGCTTCTGGAAGAGCAGACCCAAA 3597  
QY 3553 TTACAGCAGCAGATGGAAGCTTGAGAGAAATCACATTTTCCGTCGACTCAAGGACTGCAA 3612  
Db 3598 TTACAGCAGCAGATGGAAGCTTGAGAGAAATCACATTTTCCGTCGACTCAAGGACTGCAA 3657  
QY 3613 GAAGCTTAGATCGGGCTGATCTAGAGCAGAGAAAGTACTTGGAGTATCAGCTG 3672  
Db 3658 GAAGCTTAGATCGGGCTGATCTAGAGCAGAGAAAGTACTTGGAGTATCAGCTG 3717  
QY 3673 GAAAAACATTGAGTCTCTATTTCTCATGAAAAGTGAATAAGGAGGCACTATTCTCAA 3732  
Db 3718 GAAAAACATTGAGTCTCTATTTCTCATGAAAAGTGAATAAGGAGGCACTATTCTCAA 3777  
QY 3733 CAAACCAAACCTCATTTGATTTTGAAGCCAAATGGAACCACTGCTTAAAAAGAAAAAG 3792  
Db 3778 CAAACCAAACCTCATTTGATTTTGAAGCCAAATGGAACCACTGCTTAAAAAGAAAA - - 3835  
QY 3793 GGTTTATTAGTCGCGAGAAAGAGGACCTGTTTACCACACACAGGTTTCCCTGTCAGTAC 3852  
Db 3836 -----AGGTCCCTTCGAGTAC 3852  
QY 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTAGAGAAAGCC 3912  
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTAGAGAAAGCC 3912  
QY 3913 CTTCAGAGACCCGATCGAGCTCGGTCGCGCCCGGAGGAGCTGCCACCGCAAGACA 3972  
Db 3913 CTTCAGAGACCCGATCGAGCTCGGTCGCGCCCGGAGGAGCTGCCACCGCAAGACA 3972  
QY 3973 ACGGACCAACCAACCCATCCACGCGCAGCCACCGCGAGGAGCAGATCGCCATGTCGCCC 4032  
Db 3973 ACGGACCAACCAACCCATCCACGCGCAGCCACCGCGAGGAGCAGATCGCCATGTCGCCC 4032  
QY 4033 ATCGTGGGTGCGCAGAGCACAGCCCGAGTGCATGAGCTGTGCGCCCGCCATCCAGC 4092  
Db 4033 ATCGTGGGTGCGCAGAGCACAGCCCGAGTGCATGAGCTGTGCGCCCGCCATCCAGC 4092  
QY 4093 CCGAGAAAGGAGCTTTCAACTCCAGGAAATTTAGTCGGCGCTTTAAGGAAACGATGAC 4152  
Db 4093 CCGAGAAAGGAGCTTTCAACTCCAGGAAATTTAGTCGGCGCTTTAAGGAAACGATGAC 4152  
QY 4153 CACAAATATTCCTCACCGATTAACGATGAGATGGAATGGAACCATGCGAGCCACAAGTGTGCTG 4212  
Db 4153 CACAAATATTCCTCACCGATTAACGATGAGATGGAATGGAACCATGCGAGCCACAAGTGTGCTG 4212  
QY 4213 TGTCTGGATACCGTGCACCTTTGGACCGCAGGCATCCAAATGTCGGAATGTCAGTGATG 4272  
Db 4213 TGTCTGGATACCGTGCACCTTTGGACCGCAGGCATCCAAATGTCGGAATGTCAGTGATG 4272  
QY 4273 TGTACCCCAAGTGTCCAGCTGTGCGAGCCACTCGGGTTCGCTGCTGAATATGCC 4332  
Db 4273 TGTACCCCAAGTGTCCAGCTGTGCGAGCCACTCGGGTTCGCTGCTGAATATGCC 4332  
QY 4333 ACACACTTCACCGAGGCCTTCTGCGGTGACAAAATGAATCCCGCAGGTCTCCAGACCAG 4392

Db 4333 ACACACTTCACCGAGGCCTTCTGCGGTGACAAAATGAATCTCCCGAGTCTCCAGACCAAG 4392  
QY 4393 GAGCCAGCAGCAGCTTGACCTGGAGAGGTGATGAAGTGGCCCGAGGATTAACAACGA 4452  
Db 4393 GAGCCAGCAGCAGCTTGACCTGGAGAGGTGATGAAGTGGCCCGAGGATTAACAACGA 4452  
QY 4453 GGACAGCAAGGCTGGGACAGAGAGTACATTGTCTCGAGGGATCAAAAGTCTCATTTAT 4512  
Db 4453 GGACAGCAAGGCTGGGACAGAGAGTACATTGTCTCGAGGGATCAAAAGTCTCATTTAT 4512  
QY 4513 GACAAATGAAGCAGAGAGCTGACAGAGCCCGTGGAGAAATTTGAGCTGTGCTCTCC 4572  
Db 4513 GACAAATGAAGCAGAGAGCTGACAGAGCCCGTGGAGAAATTTGAGCTGTGCTCTCC 4572  
QY 4573 GAGGGGATGTATCTATTTATGCTGCGTGTGCTTCCGAACTCGCAATCAGGCCAAA 4632  
Db 4573 GAGGGGATGTATCTATTTATGCTGCGTGTGCTTCCGAACTCGCAATCAGGCCAAA 4632  
QY 4633 GCA----- 4635  
Db 4633 GCAGATGCCATACATACACTGAAGATGGAATCTCACCCGACACACCCTGCTGCGCCGGG 4692  
QY 4636 ----- 4635  
Db 4693 AGAACCTCTACTTGTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCTTA 4752  
QY 4636 -----GAAAAACAGAGCTGATGCTAAACTG 4662  
Db 4753 GAATCAGTTCTCGCAGGTGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATGCTAAACTG 4812  
QY 4663 CTTGGAACTCCCTGCTGAACTGGAGGTGATGACCGCTAGACATGAACTGCACGCTG 4722  
Db 4813 CTTGGAACTCCCTGCTGAACTGGAGGTGATGACCGCTAGACATGAACTGCACGCTG 4872  
QY 4723 CCCTTCAGTGACAGGTGTGTTGGTGGCCACCGAGGAAGGCTCTACGCCCTGAATGC 4782  
Db 4873 CCCTTCAGTGACAGGTGTGTTGGTGGCCACCGAGGAAGGCTCTACGCCCTGAATGC 4932  
QY 4783 TTGAAAACTCCCTAACCCATGTCAGAGAAATTTGGAGCAGTCTTCCAAATTTATATATC 4842  
Db 4933 TTGAAAACTCCCTAACCCATGTCAGAGAAATTTGGAGCAGTCTTCCAAATTTATATATC 4992  
QY 4843 AAGGACCTGGAGAGCTTACTCATGATAGCAGGAGAGAGCGGGCACTGTGCTTTGTGGAC 4902  
Db 4993 AAGGACCTGGAGAGCTTACTCATGATAGCAGGTGAGAGCGGGCACTGTGCTTTGTGGAC 5052  
QY 4903 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 4962  
Db 5053 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA 5112  
QY 4963 CCCAACATTTTGAAGCTGTCAGAGGTGCCACTTGTGGGGCAGGCAAGATTGAGAAC 5022  
Db 5113 CCCAACATTTTGAAGCTGTCAGAGGTGCCACTTGTGGGGCAGGCAAGATTGAGAAC 5172  
QY 5023 GGGCTCTGATCTGTGACGCAATGCCAGCAAGTCTCATTTCTCGCTACAAACAAAAC 5082  
Db 5173 GGGCTCTGATCTGTGACGCAATGCCAGCAAGTCTCATTTCTCGCTACAAACAAAAC 5232  
QY 5083 CTGAGCAATACTGCACTCGGAAAGATAGAGCCTCAGAGCCCTGAGCTGTATCCAC 5142  
Db 5233 CTGAGCAATACTGCACTCGGAAAGATAGAGCCTCAGAGCCCTGAGCTGTATCCAC 5292  
QY 5143 TTCACCAATTTACAGTATCTCTATTGGAAACCAATTAATTTCTACGAAATCGACATGAAGCAG 5202  
Db 5293 TTCACCAATTTACAGTATCTCTATTGGAAACCAATTAATTTCTACGAAATCGACATGAAGCAG 5352  
QY 5203 TACAGCTCGAGGAATCTCTGGATAGAAATGACCATTCCTTGGCAGCTGCTGTGTTGCC 5262  
Db 5353 TACAGCTCGAGGAATCTCTGGATAGAAATGACCATTCCTTGGCAGCTGCTGTGTTGCC 5412  
QY 5263 GCCTCTTCCAAACAGCTTCCCTGCTCAATCGTSCAGTGAACAGCGCAGGCGAGGAG 5322







4596 AATACAGCCAAAGCAGAGTGTCCCATATACATGTGAAGATGGAATCTCACCCGCACACCACC 4655  
4636 ----- 4635  
4656 TCCTGGCCGGGAGAACCTCTACTTGTAGCTCCAGCTCCCTGACAAACAGCGCTGG 4715  
4636 -----GAAAAAGCAGAGCT 4650  
4716 GTCAACCGCTTAGAATCAGTTGTCCAGAGTGGGAGAGTTTCTAGGGAAAAAGCAGAGCT 4775  
4651 GATGCTAAACTGCTTGGAAACTCCCTGCTGAAACTGGAAGTGTATGACCGTCTAGACATG 4710  
4776 GATGCTAAACTGCTTGGAAACTCCCTGCTGAAACTGGAAGTGTATGACCGTCTAGACATG 4835  
4711 AACTCAGCGCTGCCCTTCAGTACACAGGTGTGTGTGGGACCGAGAGAGGGCTCTTAC 4770  
4836 AACTCAGCGCTGCCCTTCAGTACACAGGTGTGTGTGGGACCGAGAGAGGGCTCTTAC 4895  
4771 GCCCTGAATGTCTGAAAACTCCCTTAACCCATGTCCTCAGGAATTTGGAGCAGTCTTCCA 4830  
4896 GCCCTGAATGTCTGAAAACTCCCTTAACCCATGTCCTCAGGAATTTGGAGCAGTCTTCCA 4955  
4831 ATTTATATATCAAGGACCTGGAGAAGTCTACTCATGATAGCAGAGAGAGCGGCACTG 4890  
4956 ATTTATATATCAAGGACCTGGAGAAGTCTACTCATGATAGCAGAGAGAGCGGCACTG 5015  
4891 TGTCTTGTGGAGTGAAGAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAG 4950  
5016 TGTCTTGTGGAGTGAAGAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAG 5075  
4951 CCGGACATCTCACCAACATTTTGAAGCTGTCAAGGGTGCACCTGTGTTGGGCGAGGC 5010  
5076 CCGGACATCTCACCAACATTTTGAAGCTGTCAAGGGTGCACCTGTGTTGGGCGAGGC 5135  
5011 AAGATTGAGAACGGGCTGTGCATCTGTGAGCCATGCCCCAGCAAAAGTGTGATTTCTCCGC 5070  
5136 AAGATTGAGAACGGGCTGTGCATCTGTGAGCCATGCCCCAGCAAAAGTGTGATTTCTCCGC 5195  
5071 TACACGAAACCTCAGCAATCTGCAATCCCGAAGAGATAGACCTTCAGAGCCCTGC 5130  
5196 TACACGAAACCTCAGCAATCTGCAATCCCGAAGAGATAGACCTTCAGAGCCCTGC 5255  
5131 AGCTGTATCCACTTCAACCAATACAGTATCTCATTTGGAACCAATAAATTTACGAAATC 5190  
5256 AGCTGTATCCACTTCAACCAATACAGTATCTCATTTGGAACCAATAAATTTACGAAATC 5315  
5191 GACATGAAGAGTACAGCTCGAGGAATCTTGGATAGAAATGACCATTCCTTGGCACT 5250  
5316 GACATGAAGAGTACAGCTCGAGGAATCTTGGATAGAAATGACCATTCCTTGGCACT 5375  
5251 GCTGTGTTGGCGCTTCTTCAACAGCTTCCCTGTCTCAATCGTGAGGTGAACAGGCA 5310  
5376 GCTGTGTTGGCGCTTCTTCAACAGCTTCCCTGTCTCAATCGTGAGGTGAACAGGCA 5435  
5311 GGGCAGCGAGGAGTACTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGATTTCTTAC 5370  
5436 GGGCAGCGAGGAGTACTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGATTTCTTAC 5495  
5371 GGAAGAGTACCGGACAGAGATCTCAAGTGGAGTCCGCTTACCTTGGCTTTGCTTAC 5430  
5496 GGAAGAGTACCGGACAGAGATCTCAAGTGGAGTCCGCTTACCTTGGCTTTGCTTAC 5555  
5431 AGAGAACCTTATCTGTTTGTGACCACTTCAACTCACTCACTCACTCACTCACTCACTCA 5490  
5556 AGAGAACCTTATCTGTTTGTGACCACTTCAACTCACTCACTCACTCACTCACTCACTCA 5615  
5491 CGCTCTCAGAGGAGCCCTGCGGAGGTACTGGACATCCGAAACCGGCTTACCTG 5550  
5616 CGCTCTCAGAGGAGCCCTGCGGAGGTACTGGACATCCGAAACCGGCTTACCTG 5675  
5551 GGCCTGCGCAATTTCTCAGAGGAGTACTTGGCGTCTCATACAGGATAAATTAAGG 5610  
5676 GGCCTGCGCAATTTCTCAGAGGAGTACTTGGCGTCTCATACAGGATAAATTAAGG 5735

5611 GTCAATTTGTCGAGGGAACCTCGTGAAGAGTCCGGCACTGAAACACACACCGGGGCCG 5670  
5736 GTCAATTTGTCGAGGGAACCTCGTGAAGAGTCCGGCACTGAAACACACCGGGGCCG 5795  
5671 TCCACCTCCCGCAGCAGCCCAACAAAGCAGAGGCCCAACCAAGTACAAACGAGCACATCACC 5730  
5796 TCCACCTCCCGCAGCAGCCCAACAAAGCAGAGGCCCAACCAAGTACAAACGAGCACATCACC 5855  
5731 AAGCGGTGCTTCCAGCCAGCGCCCGCGAAGGCCCAAGCCACCCCGGAGAGCCAAGC 5790  
5856 AAGCGGTGCTTCCAGCCAGCGCCCGCGAAGGCCCAAGCCACCCCGGAGAGCCAAGC 5915  
5791 ACACCCACCGCTACCGCGAGGGCGGACCGAGCTGCGAGGGCAAGTCTCCTGCGCCG 5850  
5916 ACACCCACCGCTACCGCGAGGGCGGACCGAGCTGCGAGGGCAAGTCTCCTGCGCCG 5975  
5851 CCCTGAGCGAGAGAGTCCCCCGCGGATCTCAGCACGCGAGAGAGCGGTCCCC 5910  
5976 CCCTGAGCGAGAGAGTCCCCCGCGGATCTCAGCACGCGAGAGAGCGGTCCCC 6035  
5911 GGGAGGTGTTTGAAGCAGCAGCAGGGCGCGCTGCTGCGGAGCGGTGAGAGCCCG 5970  
6036 GGGAGGTGTTTGAAGCAGCAGCAGGGCGCGCTGCTGCGGAGCGGTGAGAGCCCG 6095  
5971 CTGTCCAGGTGAACAAAGGGAAGAGGGCAGAGTGCCTCTCAAGTTTTCACGGTTAACACT 6030  
6096 CTGTCCAGGTGAACAAAGTCTGGGCCAGTCTCAGTATAAATCTCAGCCAGAAAAACC 6155  
6031 GTCACTTATTAT 6042  
6156 AACTCCTCATCT 6167

RESULT 12  
AX574427  
LOCUS AX574427 5877 bp DNA linear PAT 07-JAN-2003  
DEFINITION Sequence 3 from Patent WO02059325.  
ACCESSION AX574427  
VERSION AX574427.1 GI:27551752  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Yu, X., Miranda, M. and Friddle, C.J.  
TITLE Human kinases and polynucleotides encoding the same  
JOURNAL Patent: WO 02059325-A 3 01-AUG-2002;  
Lexicon Genetics Incorporated (US)  
FEATURES  
source  
1. 5877  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 87.0%; Score 5358; DB 6; Length 5877;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 5634; Conservative 0; Mismatches 5; Indels 243; Gaps 3;  
QY 1 ATGTTGAAGTCAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGTGCTGAACCCATT 60  
DB 1 ATGTTGAAGTCAATATGAGCGCGGAATCCTTTGGATGCTGTGCTGTGCTGAACCCATT 60  
QY 61 GCCAGCCGGGCTCCAGGCTGAATCTGTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120  
DB 61 GCCAGCCGGGCTCCAGGCTGAATCTGTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120  
QY 121 CAGCAGATGCTCCTTTTCCGAGAGGAGATATAGATGCCCTCTTGTCTCTTTTGA 180  
DB 121 CAGCAGATGCTCCTTTTCCGAGAGGAGATATAGATGCCCTCTTGTCTCTTTTGA 180

QY 181 GAATGCAAGTCTGCTGATGAAGATTAAAGCACTGAGCAACTTTGTCCGGAAGTAT 240  
Db 181 GAATGCAAGTCTGCTGATGAAGATTAAAGCACTGAGCAACTTTGTCCGGAAGTAT 240  
QY 241 TCCGACACCAATAGTGTAGTTACAGGAGCTCCAGCTTTCGCAAGAGACTTCGAAGTCAGA 300  
Db 241 TCCGACACCAATAGTGTAGTTACAGGAGCTCCAGCTTTCGCAAGAGACTTCGAAGTCAGA 300  
QY 301 AGTCTTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGGTGCTTAAGAGAGAAAGCAACCGGG 360  
Db 301 AGTCTTGTAGTGTGTGCTCACTTTGCTGAAGTGCAGGTGCTTAAGAGAGAAAGCAACCGGG 360  
QY 361 GACATCTATCTATGAAGTGTAGAGAGAAAGGCTTTATTTGSCC CAGAGAGAGGTTTCA 420  
Db 361 GACATCTATCTATGAAGTGTAGAGAGAAAGGCTTTATTTGSCC CAGAGAGAGGTTTCA 420  
QY 421 TTTTGTGAGAGAGCGGACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 480  
Db 421 TTTTGTGAGAGAGCGGACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 480  
QY 481 CAGTATGCCCTTTTCAGGACAAAATACCTTTTATCTGATGAGAGAAATATCAGCCTCGAGGG 540  
Db 481 CAGTATGCCCTTTTCAGGACAAAATACCTTTTATCTGATGAGAGAAATATCAGCCTCGAGGG 540  
QY 541 GACTGCTGCTCACTTTTGTATGATATGAGGACCACTTAGATGAAAACCTGATACAGTTT 600  
Db 541 GACTGCTGCTCACTTTTGTATGATATGAGGACCACTTAGATGAAAACCTGATACAGTTT 600  
QY 601 TACTAGCTGAGCTGATTTTGCTGTTTACAGCGTTTCACTGATGGGATACCTGATCGA 660  
Db 601 TACTAGCTGAGCTGATTTTGCTGTTTACAGCGTTTCACTGATGGGATACCTGATCGA 660  
QY 661 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACATCAAGCTGGTGGATTTT 720  
Db 661 GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACATCAAGCTGGTGGATTTT 720  
QY 721 GGATCTGCCGCGAAAATGAATCAAAACAGATGGTGAATGCCAAACTCCCGATTCGGACC 780  
Db 721 GGATCTGCCGCGAAAATGAATCAAAACAGATGGTGAATGCCAAACTCCCGATTCGGACC 780  
QY 781 CCGAGTTACATGCTCCTGAAAGTGTGACTGTGATGAACCGGGATGGAAAAGCCACTAC 840  
Db 781 CCGAGTTACATGCTCCTGAAAGTGTGACTGTGATGAACCGGGATGGAAAAGCCACTAC 840  
QY 841 GGCCTGCACTGTGACGTGGTGCAGTGGCGGTGATTCCTATGATGATTTATCGGAGA 900  
Db 841 GGCCTGCACTGTGACGTGGTGCAGTGGCGGTGATTCCTATGATGATTTATCGGAGA 900  
QY 901 TCCCTTCCGAGAGGAACTCTGCGAGAACCTTCAATAACATTTATGAAATTCAGCGG 960  
Db 901 TCCCTTCCGAGAGGAACTCTGCGAGAACCTTCAATAACATTTATGAAATTCAGCGG 960  
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGCAGAGTACTTTCTTGATCTGATCAAGC 1020  
Db 961 TTTTGAATTTCCAGATGACCCCAAGTGCAGAGTACTTTCTTGATCTGATTTCAAGC 1020  
QY 1021 TTGTTGTGCGGCGAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC 1080  
Db 1021 TTGTTGTGCGGCGAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC 1080  
QY 1081 TCTAAAATTTGATGTAACATTTCTGTAACCTCTCCCTCCCTTCGTTCCACCTCAAG 1140  
Db 1081 TCTAAAATTTGATGTAACATTTCTGTAACCTCTCCCTCCCTTCGTTCCACCTCAAG 1140  
QY 1141 TCTGAGATGACACCTTCAATTTTGTATGAACAGAGAGAAATTCGTTGGTTTCTCTCT 1200  
Db 1141 TCTGAGATGACACCTTCAATTTTGTATGAACAGAGAGAAATTCGTTGGTTTCTCTCT 1200  
QY 1201 CCGTCCAGCTGAGCCCTTCAGCTTCTCGGTGAAGAACTGCCGTTTGTGGGTTTTCG 1260  
Db 1201 CCGTCCAGCTGAGCCCTTCAGCTTCTCGGTGAAGAACTGCCGTTTGTGGGTTTTCG 1260  
QY 1261 TACAGCAAGGCACTGGGATTTCTTGATAGTCTGAGTCTGTTGCTCGGCTCGGACTCC 1320

Db 1261 TACAGCAAGGCACTGGGATTTCTTGATAGTCTGAGTCTGTTGCTCGGCTCGGACTCC 1320  
QY 1321 CCTGCCAAGACTAGTCTCATGGAAGAAATCTTCTCATCAAAAGCAAGAGACTCAAGAC 1380  
Db 1321 CCTGCCAAGACTAGTCTCATGGAAGAAATCTTCTCATCAAAAGCAAGAGACTCAAGAC 1380  
QY 1381 TCTCAGGACAAAGTGTCAAGATGAGCAGAAATGACCGGTTTACATCGAGAGTGTCA 1440  
Db 1381 TCTCAGGACAAAGTGTCAAGATGAGCAGAAATGACCGGTTTACATCGAGAGTGTCA 1440  
QY 1441 GAGTGGAGGCTGTGCTTAGTCAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
Db 1441 GAGTGGAGGCTGTGCTTAGTCAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
QY 1501 TCCCTTCTGAGCAGGACCTTGTACCTACATACAGAAATGCAAGCACTGCAAGTCTCCAT 1560  
Db 1501 TCCCTTCTGAGCAGGACCTTGTACCTACATACAGAAATGCAAGCACTGCAAGTCTCCAT 1560  
QY 1561 TTGAGCAAGCACGATGGAGTGTCCAGGAGGATGCAAAAGCACTGCAAGTCTCCAT 1620  
Db 1561 TTGAGCAAGCACGATGGAGTGTCCAGGAGGATGCAAAAGCACTGCAAGTCTCCAT 1620  
QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1680  
Db 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1680  
QY 1681 GTGCAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740  
Db 1681 GTGCAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740  
QY 1741 CCGAGTGTCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGTAAGAAATCAAG 1800  
Db 1741 CCGAGTGTCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGTAAGAAATCAAG 1800  
QY 1801 CGGAAACGACAGATGTACGATTAACCTTGAAGCTTAAGGATCAAGGAGAGCTGAA 1860  
Db 1801 CGGAAACGACAGATGTACGATTAACCTTGAAGCTTAAGGATCAAGGAGAGCTGAA 1860  
QY 1861 GTGAGAAATATCGAAACTGGAGAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
Db 1861 GTGAGAAATATCGAAACTGGAGAGATCAATGTCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
QY 1921 CTCCAGAGAACTGGAGAGGCT----- 1944  
Db 1921 CTCCAGAGAACTGGAGAGGCTGTAAAGCCAGCAGGAGGCCACCGAGTGTGCGAG 1980  
QY 1945 -----GCAAGGAGCGAGCGGAGGAGCTGGAGAAAGCTGCAGAACCGAGAG 1992  
Db 1945 -----GCAAGGAGCGAGCGGAGGAGCTGGAGAAAGCTGCAGAACCGAGAG 1992  
QY 1993 GATCTTCTGAAGGATCAGAAAGAGTGTGGTGAAGTGTAGGAAAGCGCCCATTTCTCG 2052  
Db 1993 GATCTTCTGAAGGATCAGAAAGAGTGTGGTGAAGTGTAGGAAAGCGCCCATTTCTCG 2052  
QY 2041 GATTTCTTGAAGGATCAGAAAGAGTGTGGTGAAGTGTAGGAAAGCGCCCATTTCTCG 2100  
Db 2041 GATTTCTTGAAGGATCAGAAAGAGTGTGGTGAAGTGTAGGAAAGCGCCCATTTCTCG 2100  
QY 2053 GAGAAACAAAGTAAAGAGATAGAGACCATGAGCGGTAGAGAAACAGAGCTGAAGGATGAC 2112  
Db 2053 GAGAAACAAAGTAAAGAGATAGAGACCATGAGCGGTAGAGAAACAGAGCTGAAGGATGAC 2112  
QY 2101 GAGAAACAAAGTAAAGAGATAGAGACCATGAGCGGTAGAGAAACAGAGCTGAAGGATGAC 2160  
Db 2101 GAGAAACAAAGTAAAGAGATAGAGACCATGAGCGGTAGAGAAACAGAGCTGAAGGATGAC 2160  
QY 2113 ATCCAGCAAAATCCCAACAGATCCAGCAGTGGCTGTATAAAATTTCTGAGCTCGAAGAG 2172  
Db 2113 ATCCAGCAAAATCCCAACAGATCCAGCAGTGGCTGTATAAAATTTCTGAGCTCGAAGAG 2172  
QY 2173 AAACATCGGAGGCGCCCAAGTCTCAGCCCGAGCCTTAGAGAGTGCACCTGAAACAGAAAG 2232  
Db 2173 AAACATCGGAGGCGCCCAAGTCTCAGCCCGAGCCTTAGAGAGTGCACCTGAAACAGAAAG 2232  
QY 2221 AAACATCGGAGGCGCCCAAGTCTCAGCCCGAGCCTTAGAGAGTGCACCTGAAACAGAAAG 2280  
Db 2221 AAACATCGGAGGCGCCCAAGTCTCAGCCCGAGCCTTAGAGAGTGCACCTGAAACAGAAAG 2280  
QY 2233 CAGCACTATGAGAAAGATTTAAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2292  
Db 2233 CAGCACTATGAGAAAGATTTAAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2292  
QY 2281 CAGCACTATGAGAAAGATTTAAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2340  
Db 2281 CAGCACTATGAGAAAGATTTAAAGTGTGGCAATCAGATAAAGAAAGACCTGGCTGAC 2340  
QY 2293 AAGGAGCACTGGAGACATGATGAGAGACACAGGAGGCGCCCATGAGAGGCGCAA 2352  
Db 2293 AAGGAGCACTGGAGACATGATGAGAGACACAGGAGGCGCCCATGAGAGGCGCAA 2352

2341 AAGGAGACACTGGAGAACATGATGACAGAGACACGAGGAGAGGCCCATGAGAAGGCGCAA 2400 Db  
2353 ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGATGATTCAGATCCAGATCCCTGGAA 2412 QY  
2401 ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGATGATTCAGATCCAGATCCCTGGAA 2460 Db  
2413 CAGAGATTGTGAACTGTCTGAAGCCAAATAAATTCAGACAAATAGCAGTCTTTTACC 2472 QY  
2461 CAGAGATTGTGAACTGTCTGAAGCCAAATAAATTCAGACAAATAGCAGTCTTTTACC 2520 Db  
2473 CAAAGGAAATGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAAATTTTAC 2532 QY  
2521 CAAAGGAAATGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAAATTTTAC 2580 Db  
2533 CTGGAGACACAGGCTCGGAACTGGAGGCCAGAACCCGAAACTGGAGGAGCAGCTGGAG 2592 QY  
2581 CTGGAGACACAGGCTCGGAACTGGAGGCCAGAACCCGAAACTGGAGGAGCAGCTGGAG 2640 Db  
2593 AAGATCAGGCAACAGACCAAGTGAAGAATCGGCTCTGGAATCTGGAGCAAGATTG 2652 QY  
2641 AAGATCAGGCAACAGACCAAGTGAAGAATCGGCTCTGGAATCTGGAGCAAGATTG 2700 Db  
2653 CGGGAGTCTAGACACGAGGAGCAGAACTGGAGCTCAAGGCCAGCTCACAG 2712 QY  
2701 CGGGAGTCTAGACACGAGGAGCAGAACTGGAGCTCAAGGCCAGCTCACAG 2760 Db  
2713 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTCACAGCTCTGAGGCTGCAGGGCG 2772 QY  
2761 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTCACAGCTCTGAGGCTGCAGGGCG 2820 Db  
2773 GGCCTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAGAGACACAGCAAGCT 2832 QY  
2821 GGCCTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAGAGACACAGCAAGCT 2880 Db  
2833 GAAGAGGAGTCCAGGACCTCAGGCGACATAGAGATGAAATCAGGCGCAATTTGATGCT 2892 QY  
2881 GAAGAGGAGTCCAGGACCTCAGGCGACATAGAGATGAAATCAGGCGCAATTTGATGCT 2940 Db  
2893 CTTCTGTAACAGTGTACTGTATTCAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2952 QY  
2941 CTTCTGTAACAGTGTACTGTATTCAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 3000 Db  
2953 GACACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3012 QY  
3001 GACACGCTGAACTCAACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060 Db  
3013 GCGGCCAACGACGAGATTGACAACTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3072 QY  
3061 GCGGCCAACGACGAGATTGACAACTGCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3120 Db  
3073 ACGGAACGAGAGATGAGCTTACCAGCCAGAGACAAACGATGAGGCTCTGAAGACCAAG 3132 QY  
3121 ACGGAACGAGAGATGAGCTTACCAGCCAGAGACAAACGATGAGGCTCTGAAGACCAAG 3180 Db  
3133 TGCACCATCTGGAGAACAGCTCATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3192 QY  
3181 TGCACCATCTGGAGAACAGCTCATGGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3240 Db  
3193 AAAGAGCGGAGTGGAGGCGCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGTTTGAAG 3252 QY  
3241 AAAGAGCGGAGTGGAGGCGCTGGAGAGCGCTCTGGGTGATGAGAAATCCAGTTTGAAG 3300 Db  
3253 TGTCTGGGTTTCGAGAGCTGAGAGATGCTGGACACCCGAGAAACAGACAGAGGCGGAGGCC 3312 QY  
3301 TGTCTGGGTTTCGAGAGCTGAGAGATGCTGGACACCCGAGAAACAGACAGAGGCGGAGGCC 3360 Db  
3313 GATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGAGTGGAGGAGCACAAGGCT 3372 QY  
3361 GATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGAGTGGAGGAGCACAAGGCT 3420 Db  
3373 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGCTGAAGGCCGAGAGCTCTCT 3432 QY  
3421 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGCTGAAGGCCGAGAGCTCTCT 3480 Db

3433 GACAAGCTCAATGACCTGGAGAACAGCATGCTATGCTTGAATGAATGCCGAGCTTA 3492 QY  
3481 GACAAGCTCAATGACCTGGAGAACAGCATGCTATGCTTGAATGAATGCCGAGCTTA 3540 Db  
3493 CAGCAGAAGCTGGAGAGCTGAACAGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAGACCAAA 3552 QY  
3541 CAGCAGAAGCTGGAGAGCTGAACAGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAGACCAAA 3600 Db  
3553 TTACAGCAGAGATGGAGCCTGACAGAAAAATCAATTTTCGCTGACCTCAAGGACTGCA 3612 QY  
3601 TTACAGCAGAGATGGAGCCTGACAGAAAAATCAATTTTCGCTGACCTCAAGGACTGCA 3660 Db  
3613 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATTTGGAGTATCAGCTG 3672 QY  
3661 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATTTGGAGTATCAGCTG 3720 Db  
3673 GAAACATTTCAAGTTCTTATTTCTGAAAGAGTGAATAATGGAAGGACATATTTCTCAA 3732 QY  
3721 GAAACATTTCAAGTTCTTATTTCTGAAAGAGTGAATAATGGAAGGACATATTTCTCAA 3780 Db  
3733 CAAACCAAACTCATTGATTTTCTGCAAGCCAAATGACCAACCTGCTTAAAGAAAAAG 3792 QY  
3781 CAAACCAAACTCATTGATTTTCTGCAAGCCAAATGACCAACCTGCTTAAAGAAAAAG -- 3838 Db  
3793 GGTTTATTTAGTCGACGAGAAAGAGGCCCTGCTTTACCCACACAGGTTCTCTCAGTAC 3852 QY  
3839 -----AGTTTCTCTCAGTAC 3855 Db  
3853 AATGAGCTGAAGCTGGCGCTGGAGAGGAGAAAGCTGCTGTGTCAGAGCTAGAGAGGCC 3912 QY  
3856 AATGAGCTGAAGCTGGCGCTGGAGAGGAGAAAGCTGCTGTGTCAGAGCTAGAGAGGCC 3915 Db  
3913 CTTCAGAAAGACCCGATCGAGCTCCGCTCCGCCCGGAGGAAGCTGCCACCGCAAGACA 3972 QY  
3916 CTTCAGAAAGACCCGATCGAGCTCCGCTCCGCCCGGAGGAAGCTGCCACCGCAAGACA 3975 Db  
3973 ACGGACCAACCAACCCATCCAGCCAGCCAGCCAGGAGGAGAGATGCGCATGCTCCGCC 4032 QY  
3976 ACGGACCAACCAACCCATCCAGCCAGCCAGCCAGGAGGAGAGATGCGCATGCTCCGCC 4035 Db  
4033 ATCGTGGGTCGCGCAGAGCACCAGCCAGTGCATGAGCTGCTGGCCCGCCCATCCAGC 4092 QY  
4036 ATCGTGGGTCGCGCAGAGCACCAGCCAGTGCATGAGCTGCTGGCCCGCCCATCCAGC 4095 Db  
4093 CGCAGAAAGGAGTCTCAACTCCAGAGGAATTTAGTCCGGCTCTTAAAGGAACGATGCAC 4152 QY  
4096 CGCAGAAAGGAGTCTCAACTCCAGAGGAATTTAGTCCGGCTCTTAAAGGAACGATGCAC 4155 Db  
4153 CACAATTTCTCACCAGATTCAGCTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4212 QY  
4156 CACAATTTCTCACCAGATTCAGCTAGGACTGAACATGCGAGCCACAAAGTGTGCTGTG 4215 Db  
4213 TGTCTGGATACCGTGCACTTTGGACGCCAGGCAATCCAAATGTCTCGAATGTCAAGTGTG 4272 QY  
4216 TGTCTGGATACCGTGCACTTTGGACGCCAGGCAATCCAAATGTCTCGAATGTCAAGTGTG 4275 Db  
4273 TGTCAACCCCAAGTCTCCAGTGTCTGCCAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4332 QY  
4276 TGTCAACCCCAAGTCTCCAGTGTCTGCCAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4335 Db  
4333 ACACATTTCAACCGAGGCTTCTGCCGTGACAAATGAATCACTCCCGAGTCTCCAGACCAAG 4392 QY  
4336 ACACATTTCAACCGAGGCTTCTGCCGTGACAAATGAATCACTCCCGAGTCTCCAGACCAAG 4395 Db  
4393 GAGCCAGCAGAGCTTGCACCTTGAAGGGTGAAGTGAAGTGGCCAGGAATAACAAACGA 4452 QY  
4396 GAGCCAGCAGAGCTTGCACCTTGAAGGGTGAAGTGAAGTGGCCAGGAATAACAAACGA 4455 Db  
4453 GGACAGCAAGGCTGGGACAGGAGTACATTTGCTGGAGGGATCAAGAGTCTCATTTAT 4512 QY  
4456 GGACAGCAAGGCTGGGACAGGAGTACATTTGCTGGAGGGATCAAGAGTCTCATTTAT 4515 Db



DLQSLVCQKRLKFEGLCCHFFPARTDWNIRNSPPFPVPTLKKDDDTNFFDEPEK  
NSWATLCVPAEPLAFSGBELPFGVFSYKALGYLRSSSVSSVSDSPAKVSMKKL  
LTKSELODSQDKHMEQEMTRLHRRVSEVAIVSOXELVKASETOXSILLODLAT  
YITESSLSKRLSEARMEVSQEDDKALQILHDIREQSKLKEIKEQEYQAOVBWRLM  
MNQLEDLVASRDLZEBESDDEKALAEFFRKANCOHKMKAKQKQPEVGEYS  
KLEKINAEQOLKIBELQKLEKRAKSTATELLQNIROKAREARELEKHNRDSS  
EGIKKLVBAERERHSHKVRLEMERRENRLKDDIQTKSEIQQWADKILLELEK  
HREAOVSAGHLNHLKQBOHYEERI KVLNDQIKKDLADKESLENMORHEEAEHKG  
KILSBQKAMINAMDSKIRSLBORIVELSEANKLAANSILFTORNKKAQEMI SELRQ  
KYLFTQAGKLAQNRKLEBOLEKISHODHSDKSLLELTRLRVLSLEHEEQKLEL  
ROLTLOLSQRESQLTALOARAALESQLEAKTELEETTAZAEERIQALTAHRE  
IORKFDALNSCTVITDEEQUNQUTEDONALNMFYLSQLDSEAGNDEIVOLRS  
BYDLHREITEREMOLTQOKTMEALKTTCTNMLEQVLDLEALNDELLEKHEQWAWR  
SVLGEKQFECRVELQRLMQLTESQARADQRI TESQVVELAVKHEKAEILALQ  
ALKEQKLAESLNDLLEKHAMLENARSLOQLLETERELKORLEEEQAKQQOM  
DLQKXHLPRITLOGLBALDRALLKTERSDLEYOLENTOVLYSHKVMWEGTISQTK  
LIDFQAKMDQAPKAPQLOVYNELKLALEKXKPCASLEALOKTRTELRSABEA  
HRKADHPHSPATARQOIANSAIVRSEHOPSPMSILLAPSSRRKSSPPEFSRR  
LKERHMHNPFRNGLNRRATCAVCLIDTVHFGRQASKLECFQWCHPKKCTCLUPAT  
CGLPAYATHFTFARCRDMNSPQLOSKEPSSSLHLEGWVPRNKXQQGWDRKYI  
VLEGSKVITDYNREAGORPVEFELCLPDGDVSIHGAVASELANAKADVPYILK  
MESHPHTTICWPQRTLYLLAPSPQKRWTALESVAVGRVSRKAEADAKLGNLSL  
KLEGGDRLDNMCTLPFSDOVLVGTTEGLYALNVLKNSLTHIPGIGAVFOIYIKOLE  
KLMLTAGERALLCLVDKVKOSLAQSHLPAQPDVSPNI FEAVGCHLPAAGKLENSL  
CLICAMPKSVILNRNDMLSKYCIKREIETSEPCSIHPTNYSILITGNKPYEIDMKQ  
YTLDFLDKNSHSLAPVAFSSNSFPVSIIVONAGORBEYLLCFHFEGVFDVSYGR  
RSRTDOLKWSRULPAYAREPVLVTFHNSLEVI EIQARSLSGSPARAYLEIPNRYL  
GPAISSGAIYASSYQDKLRLVICCKGNLVKESGTQHRVPSSTRSNKRPPTYNH  
ITKRVASSPAPPEPSHPRETPHRYDRDREGRTELRRDKSPGRPLEREKSPGRMLST  
RRERSFGRLFEDSSRGRLPAGAVRTPLSQVKNVWDQSSV"

ORIGIN

Query Match 75.1%; Score 4627; DB 10; Length 6954;  
Best Local Similarity 85.7%; Pred. No. 0;  
Matches 5332; Conservative 0; Mismatches 635; Indels 254; Gaps 7;  
QY 1 ATGTTGAAGTCAAAATATGAGCGCGGGAATCCTTTGGATGCTGGTCTGCTCAACCCCAATT 60  
DB 772 ATGTTGAAGTTCAAGTATGTTGCGGAACCCGCGGAGCGCAGTGCCTCCGAGCCCAATT 831  
QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCCCTTTATGACTCAA 120  
DB 832 GCCAGTCGGGCTCCAGGCTAAATCTCTTCTCCAGGGGAAACCCGCTCATGACTCAA 891  
QY 121 CAGCAGATGCTCTCTTTCCGAGAGAGGATATAGATGCCCTCTTTGTTCTCTTTGAA 180  
DB 892 CAGCAGATGCTGCTCTTTCCGCGGAAGGATGCTAGACGCCCTCTTGCTCTCTTTGAA 951  
QY 181 GAATGCAGTCAGCTGCTCTGATGAAGATTAAGCACGTGAGCAACTTTGTCGGGAAGTAT 240  
DB 952 GAGTGCAGCAACCCGCTGATGAGATGAAGCACGTGAGCAGCTTTGTCAGAGTAT 1011  
QY 241 TCCGACACCATAGCTGAGTTA CAGGAGCTCCAGCCTTCGGCAAGAGCTTCGAAGTGAGA 300  
DB 1012 TCCGACACCATAGCCGAGTTGGGGAGCTGCAGCCGCTCGGGCAGAGACTTCGAAGTTCGA 1071  
QY 301 AGTCATTGAGTTGCTGCTCACTTTGCTGAAGTGCAGGTGCTGAAGAGAAACCAACCCGG 360  
DB 1072 AGCTTTGUGGCTGTGCTCACTTCGCTGAAGTGCAGGTGCTGAAGAGAGCGACCCGG 1131  
QY 361 GACATCTATGCTATGAAGTGTGAAGAGAGAGGCTTTATTGCGCCAGGAGCAGGTTTCA 420  
DB 1132 GACGTCTATGCCATGAATAATCATGAAGAGAGAGGCTTTGCTGGCCAGGAAACAGGTTCA 1191  
QY 421 TTTTGTGAGAGAGCGGACATATTTATCTGAGGACAGAGCCGCTGGATCCCCAATTA 480  
DB 1192 TTTTTCGAGAGAGAGGAGACATATTTATCTCGAGCAGCAGTCTTGGATCCCCCAGTTA 1251  
QY 481 CAGTATGCTCTTTCAGGACAAAATCACTTTTATCTGATGAGGAGATATCAGCCTCGAGGG 540  
DB 1252 CAGTACGCTTTTCAGGACAAAATAACCTTTTACCTGCTCATGGAATATCAGCCTCGAGG 1311  
QY 541 GACTTCTGCTCACTTTTGAATAGATAGAGGACCGAGTTCAGTTCAGTTCAGTTCAGTTC 600

DB 1312 GATTTCCTGCTGCTTCTGAAACAGATACGAGGACCAATAGATGAGATGATCAGTTC 1371  
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGGTTTCACTGATGGGATAGTGCATCGA 660  
DB 1372 TACCTTCTGAGCTGATTTTGGCTGTTTCCAGCGGTGCACCAGATGGGATATGTCATCGA 1431  
QY 661 GACATCAAGCCTCAGAAACATTTCTGTTGACCGCAGCAGACACATCAAGCTGGTGGATTT 720  
DB 1432 GACATCAAGCCCGAGAACATCTCTCATCGACGGGAGGAGATCAAGCTGGTGGATTT 1491  
QY 721 GGATCTGCGCGGAAAATGAATTCACACAGATGGTGAATGCCAAATCTCCCGATTTGGGACC 780  
DB 1492 GGATCAGCGGTAAAGATGAATTCACAAATAA---GGTGGATCCAAACTCCCACTTTGGGACC 1548  
QY 781 CCAGATTACATGGCTCTCTGAAGTCTCACTGTGTATGAACGGGATGAAAAGGACACCTAC 840  
DB 1549 CCGGATTACATGGCTCCGGAAGTGTTCACCGTGTATGAACGAGGACCGAAGGGGCACATAC 1608  
QY 841 GGCTGGAATGTGATCTGGTGGTCTAGTGGGCGTGAATCCCTATGAGATGATTTATGGGAGA 900  
DB 1609 GGCTTGGACTGTGACTGGTGGTCTGTCTGGAGTGTGTTCCTTATGAGATGTTTATGGGAAG 1668  
QY 901 TCCCTCTCCAGAGGAACTCTCTGCCAGAACCTTCAATAACATATATGAATTTCCACGGG 960  
DB 1669 ACCCATTCACAGAGGAACTCTCCGCCGACCTTCAACACATCATGAACATTCACGGG 1728  
QY 961 TTTTGAATTTCCAGATGACCCCAAAAGTGAAGTCACTTCTTGTATCTGATTCAGATCAAAGC 1020  
DB 1729 TTTTGAAGTTTCCAGATGATGACCCCAAAAGTTAGCAGTGAAGTCTCTGATCTGCTTCAGAGT 1788  
QY 1021 TTGTTGTGGCGCAGAAAGAGAGACTGAAAGTTTGAAGGTTCTTTGCTGCCATCTTTCTTC 1080  
DB 1789 CTGCTGTGTGTCAGAAAGAGAGACTGAAAGTTTCGAGGGTCTCTGCTGCCACCTTTCTTT 1848  
QY 1081 TCTAAAAATTGACTGGAAACAACTTCGTAACCTCTCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1140  
DB 1849 GCCAGACGGACTGGAACACATCCGTAACCTCTCTCTCCCTTCGTCCTCCCAACCTCAAG 1908  
QY 1141 TCTGAGATGACACTCCAAATTTTGTATGAACACAGAGAAATTCGTGGG---TTTCATCCTC 1199  
DB 1909 TCTGACGATGACACTCCAAATTTTGTATGAACACAGAGAAATTCGTGGGCTTTTCATCCTC 1968  
QY 1200 TCCGTGCCAGTGAAGCCCTCAGCCTTCTCGGGTGAAGAACTGCCGTTTCGTGGGTTTTC 1259  
DB 1969 TGTGTGCCAGTGAAGCCCTC---GGTTCTCAGGGAAGAGCTGCCGTTTGTGGATTTTC 2027  
QY 1260 GTACAGCAAGGCACTGGGGATTTCTTGGTAGATCTGAGTCTGTTTGTGTCGGGTCTGGAATC 1319  
DB 2028 GTACAGCAAGGCACTGGGGTATCTTGGTAGATCTGAGTCTGTCGTGTCGAGTCTGGAATC 2087  
QY 1320 GCCTGCCAAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTACAGA 1379  
DB 2088 CCCTGCCAAGGTTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTCCAGA 2147  
QY 1380 CTCTCAGCAAGTGTCAAGATGGAAGAGGAAATGACCCGGTTACATCGGAGAGTGTCT 1439  
DB 2148 CTCCAGGACAGTGTCTCAAGATGGAAGAGGAAATGACCCGGTTACATCGAGAGTGTCT 2207  
QY 1440 AGAGTGGAGGCTGTGCTTGTAGTCAGAGAGGTTGGAGCTGAAGGCTCTCAGACTCAGAG 1499  
DB 2208 AGAGTGGAGGCTGTGCTTGTAGTCAGAGAGGTTGGAGCTGAAGGCTCTCAGACTCAGAG 2267  
QY 1500 ATCCCTCTCTGAGCAGGACCTTGTCTACATCAGAGATGAGTACGTTTAAAGCGAAG 1559  
DB 2268 ATCCCTCTCTGAGCAGGACCTTGTCTACATCAGAGATGAGTACGTTTAAAGCGAAG 2327  
QY 1560 TTTGGAGCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTTCCA 1619  
DB 2328 TTTGGAGCAAGCAGGATGGAGTGTCCAGGAGGATGACAAAGCTCTGAGCTTCTTCCA 2387  
QY 1620 TGATATCAGAGCAGAGCCGGAGCTCCAGAAATCAGAGCAGGAGTACAGGCTCA 1679



QY	3792	GGGTTTATTAGTCGAGGAAAGAGGACCTGCTTTACCCACACAGGTTCTCTGCGATG	3851	Db	5643	GGCCTTCAGTCACAGGTGGTGGTGGGCAAGGAGGCGCTCTACGCCCTGAATGT	5702
Db	4607	-----AGTTTCTCTGCGATG	4622				
QY	3852	CAATGAGCTGAGCTGGCCCTGAGAGAGGAGAAAGCTGCTGTCAGAGCTAGAGGAGC	3911	QY	4782	CTTGAAAACTCCCTAAACCCATGCCAGGAATGGAGCAGTCTTCCAAATTTATTTAT	4841
Db	4623	CAATGAGCTGAGCTAGCCCTGGAAGAGGAGAAAGCCGATGGCGAGCTGAGGAGGC	4682	Db	5703	CTTGAAAACTCCCTAAACCCATGCCAGGAATGGGCGAGTCTTCCAAATTTATCATCAT	5762
QY	3912	CTTTGAGAGACCCGATCGAGCTCCGCTCCGCGCGGAGGAGCTGCCACCGCAAAAGC	3971	QY	4842	CAAGGACCTGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTCTTGTGGA	4901
Db	4683	CTTTGAGAGACCCGATCGAACTCCGCTCTGCGCGGAGGAGCTGCCACCGCAAAAGC	4742	Db	5763	CAAGGACCTGAGAGCTGCTCATGATAGCAGGAGGAGAGCGGCTCTGTGCTGGTGGG	5822
QY	3972	AACGAGCCACCCACACCCATCCAGCCAGCCACCGCGAGGAGCAGATCGCCATGTCGCG	4031	QY	4902	CGTGAAGAAAGTGAACAGTCCCTGGCCCGAGTCCACCTGCTGCCAGCCCGACATCTC	4961
Db	4743	CACGAGCCACCCGACCCGCTACGCGCAGCACTGCGAGGAGCAGATCGCCATGTCGCG	4802	Db	5823	CGTGAAGAGGTGAAGCAGTCCCTGGCGCAGTCAACCTGCTGCCAGCCCGAGCTC	5882
QY	4032	CATCGTGGCTGCGCAGAGACACAGCCAGTGCATGCGATGCTGCGCCCGCCATCCAG	4091	QY	4962	ACCCAACTTTTGAAGCTGTCAAGGCTGSCACTTGTGTTGGGCGAGGCAAGATGAGAA	5021
Db	4803	CATGTGCGGTGCGCCGAGCACAGCCCGTGCATGAGCCTGTGCTGCCCCACCATCCAG	4862	Db	5883	CCCCAACTATTCGAAACCGTCAAGGCTGSCACTTGTGCTGCTGGCAGATTCGAGAA	5942
QY	4092	CCGCAAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTCTTAAGAAACGATGCA	4151	QY	5022	CGGCTCTGCACTCTGTGCGAGCATGCCAGCAAAAGTGGTCTTCTCCGCTACAAACGAAA	5081
Db	4863	CCGCAAGAAAGGAGTCTCAACTCCAGAGGAATTCAGCCGGCTCTGAAGAGCGCATGCA	4922	Db	5943	CAGCTGTGCACTCTGCGCGCTATGCCAAGCAAGTGGTCTCTCCGCTACATGACAA	6002
QY	4152	CCACAATATCTCACCGAATCAAGTAGGATGGAACATGCGAGGCCAAGATGTGCTGT	4211	QY	5082	CCTCAGCAATATCTGCAATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAAGCTGATCCA	5141
Db	4923	CCACAATATCTCACCGTCTTAATGTGGGCTGGAACATGAGAGCCACCAAGTGGCCGT	4982	Db	6003	CCTCAGCAAGTACTGCAATCCGCAAGGAGATCGAGACCTCAGAGCCCTGCAAGCTGATCCA	6062
QY	4212	GTGCTCGGATACGCTGCACTTTGGAGCGCCAGGATCCAAATGCTCGAATGTCAGGTAT	4271	QY	5142	CTTCACCAATTTACAGATCTCTCATTTGGAACCAATAAATTTACGAAATCGACATGAGCA	5201
Db	4983	GTGCTCGGATACGCTGCACTTTGGAGCGCCAGGATCCAAATGCTCGAATGTCAGGTAT	5042	Db	6063	CTTCACCAATTTACAGATCTCTCATTTGGAACCAATAAATTTATAGATCGACATGAGCA	6122
QY	4272	GTGTCACCCCAAGTGTCCACGTGTGTCGACGCCACCTGCGCTTGCCTGTGTAATATGC	4331	QY	5202	GTACACGCTCGAGGAATTCCTGGATAAGATGACCATTCCTTGGCACCTGCTGTGTTTCG	5261
Db	5043	GTGTCACCCCAATGTCACCTGCTGTCACCTGCTGTCGCTGTCGCTGTCGCTGTCGCT	5102	Db	6123	GTACACGCTTGTGATGAGTTCCTGGACAAAGAACGACCATTCCTTGGCACCTGCTGTGTTTCG	6182
QY	4332	CACACACTTACCGAGGCGCTTCTGCGTGACAAATGAACTCCCGAGTCTCCAGCAAA	4391	QY	5262	CGCCTCTTCCAAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGAGGAGCGGAGA	5321
Db	5103	CACACACTTACAGGCGCTTCTGCGTGACAAATGAACTCCCGAGGCTCCAGGCAA	5162	Db	6183	CTCCTCTGTCAAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGAGGAGCGGAGA	6242
QY	4392	GGAGCCAGCAGCAGCTTGCACCTGGAAGGCTGGATGAAGTGCAGGAAATTAACAAACG	4451	QY	5322	GGAGTACTGTGCTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAACAGCTAG	5381
Db	5163	GGAGCCTGCGCAGCAGCTTGCACCTGGAAGGCTGGATGAAGTGCAGGAAATTAACAAACG	5222	Db	6243	AGAATACCTGCTGTGCTTCCACGAATTTGGGTGTTCTGTGATTTCTTACGGAACAGCTAG	6302
QY	4452	AGGACAGCAGGCTGGGACAGAGTACATTTGCTCTGAGGGATCAAAAGTCTCATTTA	4511	QY	5382	CCGACAGAGATCTCAAGTGGAGTCCGCTTACCTTTGGCCCTTTGCTCTACAGAGAACCTTA	5441
Db	5223	GGGACAGCAGGCTGGGACAGAGTACATTTGCTCTGAGGGATCAAAAGTCTCATTTA	5282	Db	6303	CCGACAGAGATCTCAAGTGGAGTCCGCTTACCTCTGGCCCTTGGCCCTACAGAGAACCTTA	6362
QY	4512	TGACATGAAGCCAGAGAGCTGGACAGAGGCGGTGGAAGAAATTTGAGCTGTGCTTCC	4571	QY	5442	TCTGTTTGTGACCCACCTTCAACTCACCTCGAAGTAAATGAGATCCAGGACGCTCCTCAGC	5501
Db	5283	TGACATGAAGCCAGAGAGCTGGACAGAGGCGGTGGAAGAAATTTGAGCTGTGCTTCC	5342	Db	6363	TCTGTTTGTGACTCATTTCACTCCCTGGAGATCATTTAGATCCAGGCCAGATCCTCACT	6422
QY	4572	CGACGGGATGATCTATTTATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	4631	QY	5502	AGGACCCCTGCGCGAGCTACCTGGACATCCCGAACCCCGCTTACCTGGGCCCTTGCAT	5561
Db	5343	TGACGGGACGATCTATTTATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	5402	Db	6423	GGGAGCCCTGCGCGAGCATATCTGGAAATTCGAAACCCCTCGCTTACCTGGGCCCTTGCAT	6482
QY	4632	ASCA-----	4635	QY	5562	TTCTCAGGAGCGATTTTACTTTGGGTCTCATACAGGATAAATTAAGGCTCATTTGCTG	5621
Db	5403	AGCAGATGCCATACATCCTGAAGATGGATGCTCACCCACACACCACTGCTGGCCTGG	5462	Db	6483	TTCTCAGGAGCGATTTTACTTTGGGTCTCTCATACAGGATAAATTAAGGCTCATTTGCTG	6542
QY	4636	-----	4639	QY	5622	CAAGGGAACCTCGTGAAGGAGTCCGGCATCTGAACACCAACCGGGGCCCTTCCACCTCCCG	5681
Db	5463	GAGAACCTCTACTTGTAGCACCAGCTTCCCGACAGAGCGCTGGGTCAACCGCCTT	5522	Db	6543	CAAGGGAACCTCGTGAAGGAGTCCGGCATCTGAACACCAACCGGGGCCCTTCCACCTCCCG	6602
QY	4636	-----GAAAAAGAGAGTGTCTAGGGAAGGCGCGAGTGTCTAAAT	4661	QY	5682	CAGAGCCCAACAGGAGGCCACCCAGTACAGAGCAATCAACAGCGCTGGC	5741
Db	5523	AGAACTGTGTCGAGGTGGAGATTTCTAGGGAAGGCGCGAGTGTCTAAAT	5582	Db	6603	CAGAGCCCAACAGGAGGCCACCCAGTACAGAGCAATCAACAGCGCTGGC	6662
QY	4662	GTTGGAAGTCTGCTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4721	QY	5742	CTTCAGGCCAGCCCGGAGGCCCGGAGGCCCGGAGGCCCGGAGGCCCGGAGGCCCGGAGGCC	5801
Db	5583	ACTTGGAAGTCTGCTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	5642	Db	6663	CTTCAGGCCAGCCCGGAGGCCCGGAGGCCCGGAGGCCCGGAGGCCCGGAGGCCCGGAGGCC	6722
QY	4722	GCCCTTCAGTACAGGTGGTGTGGTGGGACCCAGGAAAGGCTCTACGCCCTGAATGT	4781	QY	5802	CTACC-----GCGAGGGCGGAGCCGAGTGGCGAGGACAAAGTCTCTGGCGCGCCCT	5855

Db 6723 CTACCGAGACAGAGAGGGTCGACAGAGCTGCGCAGGACAAAGTCTCCAGCGCGCCCTCT 6782  
 QY 5856 GGAGCGAGAGAGTCCCCCGCGCGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAG 5915  
 Db 6783 GGAGCGGAGAGTCCCCAGCGCGATGCTCAGCAGCTAGGAGAGCGGTCCCGAGGAG 6842  
 QY 5916 GCTGTTTGAAGCAGCAGCAGCGCGCGCTGCTCGGAGCGGTCGAGCAGCGCGCTGTC 5975  
 Db 6843 ACTGTTTGAAGCAGCAGCAGCGCGCGCTGCTCGGAGCAGTGAAGCAGCGCGCTGTC 6902  
 QY 5976 CAGGTTGAACAGGAGAGGAGCAGTGCCTCTCAAGTTT 6016  
 Db 6903 CCAGGTTAAACAGGCTGGGACCAAGTCTCAGTAAAGTCT 6943  
  
 RESULT 14  
 AF039218  
 LOCUS 5952 bp mRNA linear ROD 02-JUL-1998  
 DEFINITION Rattus norvegicus postsynaptic density protein (citron) mRNA,  
 complete cds.  
 ACCESSION AF039218  
 VERSION AF039218.1 GI:2745839  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE Rattus norvegicus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 5952)  
 ZHANG, W., APPERSON, M.L., VASQUEZ, L.E. and KENNEDY, M.B.  
 CITRON, a PSD-95-binding protein at glutamatergic synapses on  
 inhibitory neurons  
 Unpublished  
 REFERENCE 2 (bases 1 to 5952)  
 ZHANG, W., APPERSON, M.L. and KENNEDY, M.B.  
 DIRECT SUBMISSION  
 JOURNAL Submitted (18-DEC-1997) Biology, Caltech, Pasadena, CA 91125, USA  
 FEATURES  
 source  
 1..5952  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 gene  
 1..5952  
 /gene="citron"  
 CDS  
 613..5469  
 /gene="citron"  
 /function="interacting with Rho family GTPases"  
 /codon\_start=1  
 /product="postsynaptic density protein"  
 /protein\_id="AAC25483.1"  
 /db\_xref="GI:2745840"  
 /translation="MPEQMARLHRAVSEVAVLSQKVELKASETQSRSLLEODLATYI  
 TEGSKRLSEQRNVSDEDDKALQLLDHIREQSKLQEIKEQYQVEMRLMN  
 QLEEDLVARRSDLLYESLRESRLAAEFKRNKCEQHKMKAKDLGKPEVGECSRL  
 EKINAEQOLKIQELQKLEKAVKATELLNTRQAKERAERLEKLEKPEVGECSRL  
 IKKLVAEERHSLNENKVKYLETRENNRLKDDITQKSEIQQMAKXILELEKRI  
 EAQVSHLEVLHKLQKQYVEEKIYVLDNQIKDLADKESLETMRHHEEAHKGKI  
 LSPQKAMINAMDSKISLEQRIVELSEANKLAANSILFTORNKQAQEMISELQKQF  
 YLEQAKLEAONKLELEKLEKSHODHDKNRLLETRLEVSLEHEEKLEKLRQ  
 LTELQSLQERESQLTAAQAAALLESQRLQAKTELETTAAEAEIQAHTAREIO  
 RKFDAIRNCTVITDLEQNQLTEADNENLNNFYLSKQDEPAGANDEIVQLRSEV  
 DHLREITREMTQLTQKTEALTKTCMLBEQVMDLEALNDELLEKQWEARVS  
 LGKESQFCEVRELMDTKQSPADQRIETSRQVVELAVKHAELIALQOAL  
 KEQKLAESLSDLEKHAMLENARSLOQKLETERELKQRLLEQAKLQOQMDL  
 OKNHIPLTQGLQKLEADRLDLKTERDLEYLENLIQVLSHEKMEGTISQOQKLI  
 DFLQAKMDPAQKLEKLSRKKEDPALTPQVPLEYNELKLEKAKARCALEALOK  
 TRLEKSAEEAAHRAKTHPHSPATPAQIQAKSAIVRSEPHQPSAMSLIAPSSR  
 KEASTPEFBSRLKERNHNPHEFNVLGNMRAKCAVCLDTVHFGRASKCLSCQV  
 MCHPKSCCLPAKCGLPAYATHFTFAFCRDVSSPGLQSKPSSSLHLEGKMLFPRN  
 NKFGQQQWDRKYVLESGKVLIDYNEAAGRPVEEFELCLPDGDSVSHLEGVASEL  
 ANTAKADVPYILKWEHSHPTTCWPGRTLYLLAPSPFDKQRTVALESVVAGVRSEK  
 AEADAADCTSCERLPVWVEKLLNSLLKLEGDRLDMNCTLPFDQVVLVGTESGLY  
 ALNVLKNSLTHISGIVAQIYIKLEKLMTAGERALCLVDVKVKQSLAQSHLP

AQPDSNPNIPEAVKGGCHLPAAGKIEKSLICICAMPKSVILRYNDNLSEKTRKELET  
 SEFCSTHFNYSILITGKFKYFIDMKQYTLBEFLDNKNDHSLAPAVFASSTNSFVPSI  
 VQANSTQREBEYLCCFEHFGVDFVDSRRSRDLDKWSRLPLAFAPVAFSTNSFVPSI  
 LEVITQARSLGTPARAVLEIENPYLGLFAISSGAIYIASSYQDLKRVICCKGNLVK  
 ESGTEOHRVPSSTSRSPNKPPTYNBHIKRVASSPAPPEGSPHPEPTPHRYDRE  
 GRTELDRKSPGRLEKSPGRLEKSPGRLEKSPGRLEKSPGRLEKSPGRLEKSPGRLEK  
 KXWDQSSV"  
 misc\_feature  
 3052..3096  
 /gene="citron"  
 /note="may be excised by alternative splicing"  
 misc\_feature  
 4066..4110  
 /gene="citron"  
 /note="may be excised by alternative splicing"  
 ORIGIN  
 Query Match 58.5%; Score 3602.6; DB 10; Length 5952;  
 Best Local Similarity 85.7%; Pred. No. 0;  
 Matches 4171; Conservative 0; Mismatches 444; Indels 252; Gaps 5;  
 QY 1399 AAGATGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTGAGAGTGGAGGCTGTGCTT 1458  
 Db 610 ATGATGAGCAGGAAATGACCCGGTTGCTATCGGAGAGTGTGAGAGTGGAGGCTGTGCTT 669  
 QY 1459 AGTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGATCCCTCTCGAGCAGGAC 1518  
 Db 670 AGTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGATCCCTCTCGAGCAGGAC 729  
 QY 1519 CTTGCTACCTACATCAGAGATGCTAGTCTTAAGCGAGTTTGGAGCAGCAGCAGATG 1578  
 Db 730 CTTGCCACCTACATCAGAGATGCTAGTCTTAAGCGAGTTTGGAGCAGCAGCAGATG 789  
 QY 1579 GAGTGTGCCAGGAGGATGACAAAGCAGTCTGATGATTCATGATATCAGAGAGCAGAGC 1638  
 Db 790 GAGTGTCTCAGGAGGATGACAAAGCAGTCTTTCAGCTTCTCCACCATCAGAGAGCAGAGC 849  
 QY 1639 CGGAAGTCCAGAGATCAAGAGCAGGAGTACAGGCTCAAGTGAAGAGAAATGAGTTG 1698  
 Db 850 CGGAAGTCCAGGAGATCAAGAGCAGGAGTACAGGCTCAGGTGGAGGAGATGAGGCTG 909  
 QY 1699 ATGATGAATCAGTTGGAAGAGGATCTGTCTCAGCAGAGAGAGCGAGTGTCTACGAA 1758  
 Db 910 ATGATGAATCAGTGGAGAGGACCTGTCTGAGCCGAGAGACGAGGATCTCTACGAG 969  
 QY 1759 TCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGATTAAGCGGAAAGCGAGAGATGT 1818  
 Db 970 TCTGAGCTGAGGAGTCTCGGCTCGCGCGAGGATTAAGCGGAAAGCGAGAGATGT 1029  
 QY 1819 CAGCATAAATGTTGAGGCTTAGGATCAAGGAGAGCTCAAGTGGAGAGATATGCGAAA 1878  
 Db 1030 CAGCATAAATGTTGAGGCTTAGGACCTAGGAGCTTGAAGGAGATGTTCGAGAGATGTTCCAGA 1089  
 QY 1879 CTGAGAGAGATCAATGCTGAGCAGCAGTCAAGATTCAGGAGTCCAGAGAAACTGGAG 1938  
 Db 1090 CTGAGAGAGATCAATGCTGAGCAGCAGTCAAGATTCAGGAGTCCAGAGAAACTGGAG 1149  
 QY 1939 AAGGCT-----GCAAG 1950  
 Db 1150 AAGGCTGTAAAGCCACACCGAGGCCACAGAGCTCTGAGAGAACTCCGCGAGGCAAG 1209  
 QY 1951 GAGCGAGCCGAGAGGAGCTGGAGAACTGCGAGAACTGCGAGAGAGATTTCTTGAAGGATC 2010  
 Db 1210 GAGCGAGCTGAGAGGAGCTGGAGAACTGCGAGAACTGCGAGAGAGATTTCTTGAAGGATC 1269  
 QY 2011 AGAAAGAGCTGGTGGAGAGTGGAGAACTGGAGAACTGGAGAACTGGAGAACTGGAG 2070  
 Db 1270 AAAAAGAACTGGTGGAGAGTGGAGAACTGGAGAACTGGAGAACTGGAGAACTGGAG 1329  
 QY 2071 CTAGAGCCTGAGAGCTGAGAGAACTGAGAGAACTGAGAGAACTGAGAGAACTGAGAG 2130  
 Db 1330 CTAGAGCCTGAGAGCTGAGAGAACTGAGAGAACTGAGAGAACTGAGAGAACTGAGAG 1389  
 QY 2131 CAGATCCAGCAGATGGCTGATTAATTTCTGAGCTCGAAGAGAACTCGGAGGCCCAA 2190

Db 1390 CAGATCCAGCAGATGGCTGATAAAAATTTCTGGAGCTGGAGGAAAGCATCTGGGAGGCTCAG 1449  
Qy 2191 GTCTAGCCGCCACCTTAGAGCTGACCTGAAACAGAAAGAGCAGCAGCTATGAGGAAAG 2250  
Db 1450 GTCTAGCTCAACACCTTAGAAGTACACCTGAAGCAGAAGGAAACAGCAGCTATGAGGAAAG 1509  
Qy 2251 ATTAAGTGTGGCAATCAGATAAGAAAGACCTGGCTGCAAGGAGACACTGAGAAC 2310  
Db 1510 ATCAAAGTGTGGCAATCAGATAAGAAAGGACTTGGCCGCAAGGAGAGCTAGAGACC 1569  
Qy 2311 ATGATGCGAGACAGAGGAGGAGGCCATGAGAGGGCAAAATTTCTCAGCGNACAGAG 2370  
Db 1570 ATGATGCGAGGAGCAGAGAGGAGGCCACGAGAAGGCGAAGATCCTCAGCGAGCAGAG 1629  
Qy 2371 GCGATGATCAATGTATGGATTCCAGATTCAGATCCCTGGAAACAGAGGATTGTGAACTG 2430  
Db 1630 GCGATGATCAACGCAATGGATTCCAGATTCAGATCCCTGGAGCAGAGGATTGTGAGCTG 1689  
Qy 2431 TCTGAGCCCAATAACTTCAGCAAAATAGCAGTCTTTTATCCCAAGGAAACATGAGGCC 2490  
Db 1690 TCAGAAGCCCAAGAGCTTGGCCAAACAGAGTCTCTTCCACCGAGGAAATGAGGCC 1749  
Qy 2491 CAAGAAGAGATGATTCTGAATCTAGGCAACAGAAATTTTACCTGGAGACACAGGCTGGG 2550  
Db 1750 CAGGAGAGATGATCTCAGAACTCAGGCGACAGAAATTTTACCTGGAGACACAGGCCGG 1809  
Qy 2551 AAGTTGGAGGCCAGAACCGAAATCTGGAGGAGCAGCTGGAGAAATCAGCCACCAAGAC 2610  
Db 1810 AAGCTGGAGGCCAGAACCGAAATCTGGAGGAGCAACTAGAGAAATCAGCCACCAAGAT 1869  
Qy 2611 CACAGTGACAAGATCGGCTGCTGGAATCTGGAGACAGAAATCTGGGAGGCTCAGCTAGAC 2670  
Db 1870 CACAGTGACAAGATCGGCTGCTGAGCTGGAGCAAGGTTGAGAGAGGTGAGCCTGGAG 1929  
Qy 2671 CACGAGGAGCAAACTGGAGCTCAAGCGCCAGCTCAGAGCTACAGCTCAGCTCCTCGAG 2730  
Db 1930 CATGAGGAGCAAGCTGGAGCTGAAGCGGAGCTCAGGAGCTCAGCTGCTCCCTGCGAG 1989  
Qy 2731 GAGCGGAGTCAGATTGACAGCTTGAAGCTGACAGGCTGACCGGGGGCCCTGGAGAGCCAGCTT 2790  
Db 1990 GAGCTGAGTCCAGCTGACGCTTGAAGCTTGAAGCGGCGCCCGGGTGGCTTGAAGCGAGCTC 2049  
Qy 2791 CGCCAGGCGAAGACAGAGCTGGAAGAGACCAAGCAGAGAGCTGGAAGAGAGATCCAGGCA 2850  
Db 2050 CGCCAGGCGAAGAGCTGGAAGAGACCACTGAGAGAGCGGAGAGAGATCCAGGCA 2109  
Qy 2851 CTCAGCGGCAATAGAGTGAATTCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACT 2910  
Db 2110 CTCAGCGCAATCGAGATGAAATCCAGCGCAAAATTCGATGCGCTTTCGCAACAGCTGTACT 2169  
Qy 2911 GTAATCAGAGCTCGAGGAGCAGCTAAACAGCTGACCGAGGACCAAGCTGAACTCAAC 2970  
Db 2170 GTATCACCAGCTTGAGGAGCAGCTGAACAGCTTCAACGAGGACCAAGCGAGCTCAAC 2229  
Qy 2971 AACCAAACTTCTACTTGTCCAAAACAACTCGATGAGGCTTCTGGCGCCAAACGACAGATT 3030  
Db 2230 AACCAAACTTCTACTGTCCAAAACAACTCGATGAGGCTTTCGGGTGCCAAACGATGAGATT 2289  
Qy 3031 GTACAACTCGAAGTGAATGAGACCATCTCCCGCGGAGATCAGGAAACGAGAGATGAG 3090  
Db 2290 GTGAGCTCGAAGCGAGGTGAGACCATCTCCCGCGTGAATCAGGAGCGGAGATGAG 2349  
Qy 3091 CTTACAGCGCAAGCAAAACGATGAGGCTCTGAAGACCAAGCTGACCATGATGAGGAA 3150  
Db 2350 CTCACAGCGAGAGCAAAACATGAGGCTCTGAAGACGACCTGCACCATGCTGAGGAG 2409  
Qy 3151 CAGGTATGGAATTTGAGGCCCTTAAACGATGAGCTGTAGAAAAGAGCGGCGAGTGGAG 3210  
Db 2410 CAGGTATGGAATTTGAGGCCCTTAAACGATGAGCTGTAGAAAAGAGCGGCGAGTGGAG 2469  
Qy 3211 GCTGAGGAGCGCTCTGAGTGAAGAAATCCAGTTTGAAGTGTGCGGTTTCGAGAGCTG 3270  
Db 2470 GCTGAGGAGCGCTCTGCGGAGCAGAAAGTCTCAGTTTTCAGTTTTCGAGTTTCGAGAGCTA 2529

Qy 3271 CAGAGAAATGCTGGACACCGAGAAACAGAGCAGGCGGAGAGCCGATCAGCGGATCACCAG 3330  
Db 2530 CAGAGGATGCTGGACACCGAGAAACAGAGCAGGCTAGAGCTGATCAGCGGATCACCAG 2589  
Qy 3331 TCTCGCCAGTGTGTGAGCTGGAGTGAAGGACACAAAGCTGAGATTCTCGCTCTGCGAG 3390  
Db 2590 TCTCGCCAGTGTGTGAGCTGGAGTGAAGGACACAAAGCCGAGATTCTCGCGCTGCGAG 2649  
Qy 3391 CAGGCTCTCAAAAGACAGAAAGCTGAAAGGCCGAGAGCTCTCTGACAAAGCTCAATGACCTG 3450  
Db 2650 CAGGCTCTCAAAAGACAGAAAGCTGAAAGCTGAGAGCTCTCGGACAAAGCTCAAGACCTG 2709  
Qy 3451 GAGAAGAGCATGCTATGCTTGAATGAATGCTCCGAAAGCTTACAGCAAGAGCTGGAGACT 3510  
Db 2710 GAGAAGAGCAGCCATGCTTGAATGAATGCTCCGAAAGCTTACAGCAAGAGCTGGAGACA 2769  
Qy 3511 GAGAGAGCTCAAAACAGAGGCTTCTGGAGAGCAAGCCAAATTTACAGCAGCAGATGAGAC 3570  
Db 2770 GAGCGGAGCTCAAAACAGAGGCTTCTGGAGAGCAAGCCAAAGTTGACGAGCAGATGAGAC 2829  
Qy 3571 CTGCAGAAAAATCACATTTTCCGTCTGACTCAAGGACTGCAAGAGCTCTAGATCGGGCT 3630  
Db 2830 CTGCAGAAAGAACACATCTTTCAGACTGACAAAGGCTGCAAGGAGGCTGAGCCCGGGC 2889  
Qy 3631 GATCTACTGAGACAGAAAGAGTGTGAGTATCAGCTGGAAGAAACATTCAGGTTCTC 3690  
Db 2890 GATCTGCTGAAGCCGAAAGAGGAGCCTGGAATATCAGCTGGAAGAAACATTCAGGTTCTC 2949  
Qy 3691 TATTCTCATGAAAGGTTGAAATGGAAGGCACTATTCTCAACAAACCAAACTCAATTGAT 3750  
Db 2950 TACTCACACAGAAAGTGAATGGAAGGCACTATTCTCAACAAACCAAACTCAATTGAT 3009  
Qy 3751 TTTCTGCAAGCCAAAATGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3810  
Db 3010 TTTCTGCAAGCCAAAATGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3069  
Qy 3811 AAGAGACACCTGCTTTTACCCACACAGGTTCTTCTGCAATGATGATGATGATGATGATG 3870  
Db 3070 AAGAGACACCTGCTTTTACCCACACAGGTTCTTCTGCAATGATGATGATGATGATG 3129  
Qy 3871 CTGCAGAAAGAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3930  
Db 3130 CTGGAAGAGAAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3189  
Qy 3931 GAGCTCCGCTCCGCGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3990  
Db 3190 GAACTCCGCTCCGCGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3249  
Qy 3991 TCCACGCGCAGCCACCGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4050  
Db 3250 TCTACGCGCAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3309  
Qy 4051 CACGAGCCAGTGGCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4110  
Db 3310 CACGAGCCAGTGGCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3369  
Qy 4111 ACTCCAGAGAAATTTAGTCCGCGCTTTAAGGAACGATGCAACCAATATTCCTCAGCGA 4170  
Db 3370 ACTCCAGAGAAATTTAGTCCGCGCTTTAAGGAACGATGCAACCAATATTCCTCAGCGG 3429  
Qy 4171 TTCAACCTTAGAGTGAACATGCGAGGACCAAAAGTGTGCTGTGTGTGTGTGTGTGTGTGTG 4230  
Db 3430 TTTAACCTTAGAGTGAACATGCGAGGACCAAAAGTGTGCTGTGTGTGTGTGTGTGTGTGTG 3489  
Qy 4231 TTTTGGAGCCAGAGCTCCAAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATG 4290  
Db 3490 TTTTGGAGCCAGAGCTCCAAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATG 3549  
Qy 4291 ACCTGCTTGGCAGCAGCTGCGGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4350  
Db 3550 ACCTGCTTGGCAGCAGCTGCGGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3609



8543060  
2 (bases 1 to 4967)  
REFERENCE  
AUTHORS  
Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuoka, T., Bito, H.,  
Ishizaki, T., and Narumiya, S.  
TITLE  
Role of Citron kinase as a target of the small GTPase Rho in  
cytokinesis  
JOURNAL  
Nature (1998) In press  
3 (bases 1 to 4967)  
REFERENCE  
AUTHORS  
Madaule, P., Eda, M., Watanabe, N., Fujisawa, K., Matsuoka, T., Bito, H.,  
Ishizaki, T., and Narumiya, S.  
TITLE  
Direct Submision  
JOURNAL  
Submitted (04-JUN-1998) Pharmacology, Faculty of Medicine, Kyoto  
University, Yoshida-Konoe, Sakyo-ku, Kyoto 606 8315, Japan  
FEATURES  
Location/Qualifiers  
1..4967  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="brain"  
/dev\_stage="adult"  
41..4928  
/function="binds to Rho and regulates cytokinesis"  
/note="kinase variant form of Citron; approximately the  
first 50 amino acids belong to the kinase domain;  
alternatively spliced form of Citron, encoded by GenBank  
Accession Number AF039218; similar to Rattus norvegicus  
Citron-K encoded by GenBank Accession Number AF070065"  
/codon\_start=3  
/product="Citron-K kinase"  
/protein\_id="AAC27933.1"  
/db\_xref="GI:3360514"  
/translation="PFVPTLKSDDDTNFDPEKNSWSSVQCLSPSGFSEBELPFV  
GFSYKALGYLRSVSSVSLDPAKVSMEKLLIKSELQDSQDKHKEQEMTDL  
HRRVSEAVLSQKVELKASETORSLLEQLATVITCSLSLKSLSEARVEVSED  
KALQQLHDIREGSRKLQIKQBYQYQVQVEKRLMNLQLEELVARRRSDLYSELRE  
SRLAAEFKRNANCOHKLMAKQDQPEVGEYSKLEKINAEQKQLEQLKLEKAV  
KASTATELLQIRQAEARELEKLNHEDSEGIKKLVEAEELVEHEKLEKQVSA  
QHEVLHKKQHEVKEIKVLNDQIKDLADKESLENNMQRHEHEKSKILSEKQA  
MINAMDSKRSRQIVELSEANKLAANSLLPTQNMKAQEMISELRQKFLYEQTA  
KLEAQNKLRLQLEKLSHQHSDKSLLELRLREVSLEHEQLEKQLTLEQL  
SQERESQITALQARALESQLRQAKTELETTAEABEETQALTAHDEIQRKEDAL  
RNSCTVITDEQLNQLTDENALNNQNFYLSKQDSEASGANDLIVQLRSVDHLRE  
ITERMQTSQKQMEALKTQTMLEEQVLDLEALNDELLEKEROWEAWRSVLGDEKS  
QPEVRVRELQMDTEKQSRADQITETESQVVELAVKEHAEILALQKALKEQKL  
AESLQKLNDELKHAMLEMAKSLQKLETERELKQRLLEEQKALQOQMDLKNHIF  
RLTQGLQALDRDLKTERSLVQLENIQVLYSHEKVMGEGTISQOTKLIDFLQAK  
MDQAKKKVLPQVNLKLALEKXKARCAELEALQKTRIELRSARESEAAHKKADHP  
HSTPATRQITAMSAIVRSPKQPSMSLLAPPSRRKESSTPEFSRRLKERMHN  
IPHRNVGLNNKACVCLDTVHFGQASKRCLEQVCHPKCSTCLPATCGLPAEYA  
THTEAFCDKNSPGLQSKPEGSLHLEGKWKVPPNNKQGGMDRYIIVLESKVL  
IYDNEARAGQRPVEFEFLCLPDGDSVHIGAVGASELANTAKADVPYILKMESPHIT  
CWPGRTLVLALSPFDKQKRWTALESVAVGVRSEKAEADAKLGNLSLLKLEGDRL  
DMNCTLPDSQVVLVGTREGLVALNVLKNSLTHIPGIGAVFQIVYIKOLEKLIMAGE  
ERALLVDVKVKQSLAQSHLPAQPDVSPNIFPAVKGCHLFAAGKIENSLCICAMPS  
KVYILRYNDNLSKICIRKEIETSEFSCSIHFTNYSILIGTNKPFYELDMKQYTLDEFLD  
KNDHSLAPAVFSSNSPEFVSIVQNSAGQREYLLCFHEFGVFDVSGRRSRDLDL  
WRLPLAFYQKLPYLVFTHFNLSLEVIEIQARSSLSGPARYLEIPNRYLHPLFAISSGA  
IYLLASSYQDKLVYICCKGNLVRESQTEHREVESTSRSPKRPPTNYEHTFKRVASS  
PAPPSPHSPRPSTPHRYRDREGRTELRRDKSPGRPLEREKSPGRMLSTRERSPGR  
LFEDSSRGRPLPAGAVRTPLSQNVKWDQSSV"

Query Match 58.4%; Score 3593.8; DB 10; Length 4967;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 4249; Conservative 0; Mismatches 527; Indels 279; Gaps 6;  
QY 1118 CCCCCTTCGTTCCACCCCTCAAGCTGACGATGACACCTCCCAATTTTGATGAACACAGAGA 1177  
DB 1 CCCCCTTCGTTCCACCCCTCAAGCTGACGATGACACCTCCCAATTTTGATGAACACAGAGA 60  
QY 1178 AGAATTCGTGGGTTTCATCTCTCCGTGCGCAGCTGAGCCCTTCAGGCTTCTCGGGTGAAG 1237  
DB 61 AGAATTCGTGGGTTTCATCTCTCTGTGTGCCAGCTGAGCCCTTCAGGCGGAAG 120

ORIGIN

QY 1238 AACTGCCGTTTGTGGGTTTTCGTACAGCAAGCAGCTGGGATTCTTGGTAGATCTGAGT 1297  
DB 121 AGTGCCTGTTGGGATTTTCGTACAGCAAGCAGCTGGGTTAATCTTGGTAGATCTGAGT 180  
QY 1298 CTGTGTGTCTGGGCTCTGGACTCCCTCCCAAGACTAGCTCCATGGAAAAAAACTTCTCA 1357  
DB 181 CTGTGTGTCTGGACTCTGGACTCCCTCCCAAGCTTAGCTCCATGGAAAAAAACTTCTCA 240  
QY 1358 TCAAAAGCAAGAGCTTACAAGACTCTCAGGACAAGTGTACAGATGGAGCAGGAATGA 1417  
DB 241 TCAAAAGCAAGAGCTTCCAAGACTCCCAAGACTCCCAAGAGTGGAGCAGGAATGA 300  
QY 1418 CCGGTTTACATCGGAGTGTACAGGTGGAGGCTGTGCTTGTAGTCAGAGGAGGTGGAGC 1477  
DB 301 CCGGTTTACATCGGAGTGTACAGGTGGAGGCTGTGCTTGTAGTCAGAGGAGGTGGAGC 360  
QY 1478 TGAAGGCTCTGAGACTCAGAGATCCCTCTCTGGAGCAGGACCTTGTCTACCTACATCAG 1537  
DB 361 TGAAGGCTCTGAGACTCAGAGATCCCTCTCTGGAGCAGGACCTTGTCTACCTACATCAG 420  
QY 1538 AATGCAGTAGCTTAAAGCGAAGTTTGAGCAACACCGGATGGAGGTGTCCCGAGGAGT 1597  
DB 421 AATGCAGTAGCTTAAAGCGAAGTTTGAGCAACACCGGATGGAGGTGTCCCGAGGAGT 480  
QY 1598 ACAAGCACTGCAGCTTCTCCATCATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCA 1657  
DB 481 ACAAGCTCTGCAGCTTCTCCAGCATCCAGAGCAGAGCCGGAAGCTCCAGGAGATCA 540  
QY 1658 AAGAGCAGGAGTACCAGGCTCAAGTGGAGAAATAGGTTGATGATGAATCAGTTGGAAG 1717  
DB 541 AGGAGCAGGAGTACCAGGCTCAGGTGGAGAGATGAGGCTGATGATGAATCAGTTGGAAG 600  
QY 1718 AGGATCTGTGTCTCAGCAAGAGCAGGAGTGTCTCTACGAATCTGAGCTCAGAGAGTCTC 1777  
DB 601 AAGACCTTGTGTCTCAGCCCGCAGACGCGAGCTCTCTACGAGTCTGAGCTGAGGAGTCTC 660  
QY 1778 GGCTTGTGTCTGAAGAAATTCAGCGGAAACGACAGATGTGACATTAATCTGTTGAAG 1837  
DB 661 GGCTTGTGTCTGAAGAAATTCAGCGGAAACGAAATTCAGTCTGACCAAACTGATGAAG 720  
QY 1838 CTAAAGATCAAGGAGCCTGAAAGTGGAGAAATATGCAAACTGGAGAAATCAATCTGCT 1897  
DB 721 CTAAGGACCAAGGAGCCTGAAAGTGGAGAAATATTCAAACTGGAGAAATCAATCTGCT 780  
QY 1898 AGCAGCAGCTCAAAATTCAGGAGCTCCAAAGAAAATGGAGAAAGGCTGCAAGAGGAGG 1957  
DB 781 AGCAGCAGCTGAAGATCCAGAGCTCCAGGAGAGCTGGAGAAAGGCTGTAAAAGCCAGCA 840  
QY 1958 CCGAGAGGAGCTGGAGAGCTGCAGAACCGAGAGGATTTCTTCTGAGGCGATCAGAAAGA 2017  
DB 841 CAGAGGCCACCGAGCTCCTCAGAAATATCCGCCAGG-----CAAAGG 882  
QY 2018 AGCTGGTGAAGCTGAGAAACCGCCGCAATTTCTCTGGAGAAACAAGGTAAGAGACTAGAGA 2077  
DB 883 AGCGAGCAGAGCGGGAG-----CTAGAGAGACTACACAACCGGGAAGACT 927  
QY 2078 CCATGGAGCGTAGAGAAAACAGACTGAAGATGACATCCAGACAAAATCCCAACAGATCC 2137  
DB 928 CCTCCGAGGCGATCAAAAAGAAAGCTGTGTGAAGCC----- 962  
QY 2138 AGCAGATGGCTGATAAAAATTTCTGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAG 2197  
DB 963 -----GAGGAGCTGGAGGAGAAACATCGGGAGGCTCAGGCTCTCAG 1002  
QY 2198 CCCAGCACCTTAGAAGTGCACCTGAAAACAGAAAAGAGCAGCATATGAGGAAAAGATTAAAG 2257  
DB 1003 CTCACATCTAGAAAGTACACTTGAACAGAAAGAACACGACTACGAGGAAAAGATCAAGG 1062  
QY 2258 TGTGGCAATCAGATAAAGAGACCTGCTGACAGGAGACACTCGAGAACATGATGC 2317  
DB 1063 TATTGGCAATCAGATAAAGAGACCTGCTGCGCGCAAGGAGAGCCTGGAGAAACATGATGC 1122

QY 2318 AGGACACGAGGAGGAGCCCATGAGAGGCGCAAAATTTCTCAGGACACAGAGCGATGA 2377  
Db 1123 AGAGACACGAGGAGGAGGCCCCAGAGAGGCGAGATCCTCAGCGAGACAGAGGCGATGA 1182  
QY 2378 TCAATGCTATGGATTCAGATCAGATCCCTGTGAAACAGAGAGATTTGTGAACTGTCTGAAG 2437  
Db 1183 TCAACGCGATGGATTCCAGATCCGATCCCTGGAGCAGAGATCGTGGAGCTGTGGAAG 1242  
QY 2438 CCATATAACTTGCAGCAAAATAGCAGTCTTTTACCCHAAAGGAACATGAAGGCCCAAGAG 2497  
Db 1243 CCACAAGCTTGGCGGCAAAACAGCAGTCTCTTACCAGAGAGAAACATGAAGGCCCAGGAAG 1302  
QY 2498 AGATGATTTCTGAACCTCAGGCAACAGAAAAATTTTACCTTGGAGACACAGGCTGGGAAGTTGG 2557  
Db 1303 AGATGATCTCAGAACTCAGGACAGCAAAATTTTACCTTGGAGACAGGCGCGAAGCTGG 1362  
QY 2558 AGSCCCAGAACCGAAACTCGAGAGCAGCTGGAGAGATCAGCCCAAGACCAAGTG 2617  
Db 1363 AGSCCCAGAACCGAAAGCTGGAAGCAACTGGAGAGATCAGCCCAAGATCAAGTG 1422  
QY 2618 ACAAGAACTGGAGCTGAAGCTGGAACAAAGTTGCGGGAGGTCACTAGAGCAGAGG 2677  
Db 1423 ACAAGAGCTGGCTGCTGGAGCTGGAACAAAGCTGAGGAGGTCACTGGAGCAGAGG 1482  
QY 2678 AGCAGAACTGGAGCTGAAGCTGGAACCAAGCTCAAGAGCTCAGCTCTCCCTGAGGAGCGCG 2737  
Db 1483 AGCAGAACTGGAGCTGAAGCTGGAACCAAGCTCAAGAGCTGCTCCCTGAGGAGCGCG 1542  
QY 2738 AGTCAGAGTTGACAGCCCTCAGAGCTGACGGCGCGCTTGGAGAGCAGCTTCGCCAGG 2797  
Db 1543 AGTCAGAGTTGACAGCCCTCAGAGCTGACGGCGCGCTTGGAGAGCAGCTTCGCCAGG 1602  
QY 2798 CGAAGCAGAGCTGGAAGAGACACAGCAGCAAGCTGAAGAGGAGATCCAGGCACTCACGG 2857  
Db 1603 CGAAGCAGAGCTGGAAGAGACACACCGGAGGAGGAGATCCAGGCGCTCACGG 1662  
QY 2858 CACATAGAGTGAATCCAGCGCAAAATTTGATGCTCTTCTGTAAGCTGTACTGTAAATCA 2917  
Db 1663 CACATCGCATGAGATCCAGCGCAAAATTTGATGCTCTTCTGTAAGCTGTACTGTAAATCA 1722  
QY 2918 CAGACCTGAGGAGCAGCTAAACACAGCTCAGCAGGACAAACCTGAATCAACAAACCAAA 2977  
Db 1723 CCGACCTGAGGAGCAGCTGAACAGCTCAGCAGGACAAACCGCGAGCTCAACAAACCAAA 1782  
QY 2978 ACTTCTACTTTGCCAAACAACTCGATGAGCTTTCTGGGCGCAACAGCAGATTTGTACAAC 3037  
Db 1783 ACTTCTACTTTGCCAAACAACTCGATGAGCTTTCTGGGCGCAACAGCAGATTTGTACAGC 1842  
QY 3038 TCGAAGTGAAGTGGACCATCTCCGCGGAGATCAGGAAAGAGAGATCGAGCTTACCA 3097  
Db 1843 TCGGAGTGAAGTGGACCATCTCCGCGGAGATCAGGAAAGAGAGATCGAGCTTACCA 1902  
QY 3098 GCCAGAACCAACGATGGAGGCTCTGAAGACCACTGACCACTGCTGGAGGAAACAGGTCA 3157  
Db 1903 GCCAGAACCAACGATGGAGGCTCTGAAGACCACTGACCACTGCTGGAGGACAGGTCC 1962  
QY 3158 TGGATTTGAGGCCCTTAACGATGAGCTGTAGAAAAAGAGCGGAGTGGAGGCTTGG 3217  
Db 1963 TGGACCTGGAGGCCCTTGAACGACGAGCTGTGGAGAAAGAGCGGAGTGGAGGCTTGGC 2022  
QY 3218 GGAGCGTCTCGGTGATGAGAAATCCAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAG 3277  
Db 2023 GGAGCGTCTTGGCAGCAGAGATCCAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAG 2082  
QY 3278 TGCTGGACACCGAGAAACAGACCGGCGAGAGCGGATCAGGGATCAGCGGCTTGGC 3337  
Db 2083 TGCTGGACACCGAGAAACAGACCGGCGAGAGCGGATCAGGGATCAGCGGCTTGGC 2142  
QY 3338 AGGTGGTGGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTTCTGCTCTGAGCAGGCTC 3397  
Db 2143 AGGTGGTGGAGTGGCGGTGAAGGAGCAACAGGCGGAGATTTCTGCTCTGAGCAGGCTC 2202  
QY 3398 TCAAGAGCAGAGCTGAAGCGCGAGAGCGCTCTCTGACAAAGCTCAATGAACCTGGAGAAGA 3457

Db 2203 TCAAGGACAGAAAGCTCAAGCCGAGAGCTGTGGACAAAGCTCAACGACCTGGAGAAGA 2262  
QY 3458 AGCATGCTATGCTTGAATGAATGCCGAAGCTTACAGCAGAAGCTGGAGACTGAACGAG 3517  
Db 2263 AGCAGCCATGCTGGAGATGAACGCCCGGAGCTTACAGCAGAACTAGACAGACAGCGGG 2322  
QY 3518 AGCTCAACACAGAGGCTTGGAGAGCAAGCCAAATTTACAGCAGCAGATGAACCTGAGA 3577  
Db 2323 AGCTCAACACAGAGGCTTGGAGGAGCAAGCCAAATTTGCAGCAGCAGATGGACCTGAGA 2382  
QY 3578 AAAATCAATTTCCGCTGACTCAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTAC 3637  
Db 2383 AGAACCACTCTCAGACTGACGCAAGGCTGACGAGGCGCTGGACCGGGCGATCTGC 2442  
QY 3638 TGAAGACAGAAAGAGTGAATTTGAGTATCAGCTTGGAAAAATTTCAAGTTTCTCTATTCTC 3697  
Db 2443 TGAAGACAGAAAGAGGCGACCTGGAGTACCAGCTTGGAAAAATTTCAAGTTTCTCTATTCTC 2502  
QY 3698 ATGAAAAGGTGAATTTGAAAGGCACTATTTCTCAACAAACCAACTCAATTTGATTTCTGC 3757  
Db 2503 ACAGAAAGTGAATTTGAAAGGCACTATTTCTCAGCAACCAACTCAATTTGATTTCTGC 2562  
QY 3758 AAGCCAAATGGACCAACCTGCTAAAAAGAAAAAGGGTTTATTTAGTCGACGGAAGAGG 3817  
Db 2563 AAGCCAAATGGACCAACCTGCTAAAAAGAAAA----- 2595  
QY 3818 ACCCTGCTTTACCAACACAGAGTTCTCTGCACTACATGAGCTGAAGCTGGCCCTGGAGA 3877  
Db 2596 -----AGTTTCTCTGCAGTACAAATGAGCTGAAGCTGGCCCTGGAAA 2637  
QY 3878 AGGAAAAGCTCGCTGTCAGAGCTAGAGGAAGCCCTTCAGAAAGCCCGCATCGAGCTCC 3937  
Db 2638 AGGAAAAGCCCGATGCGCGAGCTGGAGGAGCCCTTCAGAAAGCCCGCATCGAAGCTCC 2697  
QY 3938 GGTCCGCGGGAGGAAGCTGCCCCACCGAAAGCAACGGAACCAACCCACCCATCAACGC 3997  
Db 2698 GCTGCGCGGGAGGAAGCTGCCCCACCGAAAGCAACGGAACCAACCCGCTCTACGC 2757  
QY 3998 CAGCCACCGGAGCAGCAGATCGCCATGTCGCCATCGTCGGTGGCGGCTGGAGCAGCAGC 4057  
Db 2758 CAGCCATGCGAGCAGCAGATCGCCATGTCGCCATGTCGGTGGCGGCTGGAGCAGCAGC 2817  
QY 4058 CCAAGTGCATGAGCCTGCTGGCCCCCGCCATCCAGCCGCAAGAAAGGAGTCTTCAACTCCAG 4117  
Db 2818 CCAAGTGCATGAGCCTGCTGGCCCCCGCCATCCAGCCGCAAGAAAGGAGTCTTCAACTCCAG 2877  
QY 4118 AGGAATTTAGTCGGCGCTTAAAGAAAGCATGACCAACATATTTCTCACCAGATTCACAG 4177  
Db 2878 AGGAATTTAGTCGGCGCTTAAAGAAAGCATGACCAACATATTTCTCACCAGATTCACAG 2937  
QY 4178 TAGGACTGAACATCGAGCCACAAAGTGTGCTGTCTGGATACCGTGCATTTGGAC 4237  
Db 2938 TGGGCTTGAACATGAGAGCCACCAAGTGGCGCTGTCTGGATCTGTGCACTTTGGAC 2997  
QY 4238 GCAGGATCCAAATGTCTGAAATGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4297  
Db 2998 GCAGGATCCAAATGTCTGAAATGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3057  
QY 4298 TGGCAGCCACTCGCGCTTGCCTGCTGAATATGSCACACACTTTCACGAGGCTTCTGCC 4357  
Db 3058 TGGCCTTACTTGTGGCTTGCAGCTGATATGSCACACACTTTCACGAGGCTTCTGCC 3117  
QY 4358 GTGACAAAATGAATCTCCAGAGTCTCCAGCAAGAGGAGCCAGCAGAGCTTGGACCTGG 4417  
Db 3118 GGGACAAAATGAATCTCCCGGGCTCCAGAGCAAGAGGCTTGGCAGAGCTTGGACCTGG 3177  
QY 4418 AAGGTGGATGAAGTGGCCAGCAATTAACAAACGAGGACAGCAAGCTGGGACAGGAAGT 4477  
Db 3178 AAGGTGGATGAAGTGGCCAGCAATTAACAAACGAGGACAGCAAGCTGGGACAGGAAGT 3237  
QY 4478 ACATTTCTCTGGAGGATCAAAAGTCTCTATTTATGACAAATGAACCCAGAGAGCTGGAC 4537

```
Db 3238 ACATTGTTCTGGAGGGGTCAAAAGTCTCTCATCTATGACATGAAGCCAGAGAAGCTGGAC 3297
Qy 4538 AGAGCCCGGTGGAGAAATTTGAGCTGTGCTTCCGACGGGGATGTATCTATCTCATGGTG 4597
Db 3298 AGAGCCCGGTGGAGAAATTTGAGCTGTGCTTCCGACGGGGATGTATCTATCTCATGGTG 3357
Qy 4598 CCGTTGTTGCTTCCGAAGTCTGCAAAATACAGCCAAAGCA----- 4635
Db 3358 CCGTTGTTGCTTCAAGAACTTGCANAATACCGCAAGCAGATGTCCCATACATCTCGAAGA 3417
Qy 4636 ----- 4635
Db 3418 TGGAGTCTCACCCACACACACCACCTGCTGGCCTGGGAGAACCTCTACTTGTAGCAOCCA 3477
Qy 4636 ----- 4635
Db 3478 GCTTCCCCGCAACAGCAGCGCTGGGTCAACGGCTTAGAATCTGTGTCGTCGACAGTGGGAGAG 3537
Qy 4636 -----GAAAAAGCAGAAGTGTATGCTAAACTGCTTTGSAAACTCCCTGCTGTAACATGG 4687
Db 3538 TTCTTAGGMAAAGCCGAAGCCGATGCTAAATTAATTTGGAATCTCTCTGCTGAAACTGG 3597
Qy 4688 AAGTGTATGACCGTCTAGACATGAATGCACTGCACTGCTTCCAGTGACCAAGTGTGTGG 4747
Db 3598 AGGGCGATGACCGGCTTGACATGAATGCACTGCACTGCTTCCAGTGACCAAGTGTGTGG 3657
Qy 4748 TGGGCACCGAGGAAGGGTCTACGCGCTGAATGTCTTGAAAACTTCCCTAAACCCATGTCC 4807
Db 3658 TGGGCACCGAGGAAGGGCTCTACGCGCTGAATGTCTTGAAAACTTCCCTAAACCCATGTCC 3717
Qy 4808 CAGGAATTTGAGCAGTCTTCAAATTTATATATTAATTAATCAAGGAACTGGAGAACTCATGA 4867
Db 3718 CAGGAATTTGAGCAGTCTTCAAATTTATATATTAATTAATCAAGGAACTGGAGAACTCATGA 3777
Qy 4868 TAGCAGGAAGCAGCGGCACTGTCTTGTGAGCTGAAGAAAGTGAACAGTCCCTGG 4927
Db 3778 TAGCAGGAAGCAGCGGCTGTCTGCTGCTGAGCTGAAGAAAGTGAAGCAGTCCCTGG 3837
Qy 4928 CCCAGTCCCACTGCTGCTGCCAGCCGACATCTCACCAACATTTTGAAGCTGCAAG 4987
Db 3838 CGCAGTCAACCTGCTGCTGCCAGCCGACGCTCTCCCCAACATATTCGAAGCGCTCAAG 3897
Qy 4988 GCTGCCACTTTTGGGGCAGCAGATGTAAGAAAGGCTCTGCACTGTGAGCAGCATGC 5047
Db 3898 GCTGCCACTTTGCTGCTGTCAGATCGAAGCAGCTGTGCACTGTGCGCCGCTATGC 3957
Qy 5048 CCAGCAAGCTGCTCATTTCTCCGCTACACGAAACCTCAGCAAAATACTGCATCCGGAAG 5107
Db 3958 CAAGCAAGTGTCTATCTCCGCTACATGAACAACCTCAGCAAGTACTGCATCCGCAAG 4017
Qy 5108 AGATAGAGACCTCAGAGCCCTGAGCTGTATCCACTTCAACAAATTAAGTATCTCATGG 5167
Db 4018 AGATCAGAGACCTCAGAGCCCTGAGCTGTATCCACTTCAACAAATTAAGTATCTCATGG 4077
Qy 5168 GAACCAATTAATCTACGAATTCGACATGACAGCAGTACAGCTCGAGGAATTCCTGGATA 5227
Db 4078 GGACCAACAAATTTCTATGAGATCGCATGAGCAGTACAGCTTGTATGATGATTCCTGGACA 4137
Qy 5228 AGAATGACCAATTCCTTGGACCTGCTGTGTTGTCGCGCTCTTCCAAACAGCTTCCCTGTCT 5287
Db 4138 AGAACGACCATTCCTTTGGCACCTGCTGTGTTGCTCTCTCGTCCAACAGCTTCCCTGTCT 4197
Qy 5288 CAATGCTGAGGTGAACAGCGCAGGAGGAGGAGTACTGTCTGTGTTTCCAGCAAT 5347
Db 4198 CCATTGTGACGGCAACAGCGCGGGCAGGAGAAATACCTGTGTGCTTCCAGCAAT 4257
Qy 5348 TTGGAGTGTCTGTTGATTTTACGGAAGAGCTAGCCGCAAGACGATCTCAAGTGGAGTC 5407
Db 4258 TTGGGTGTTCTGTTGATTTCTTACGGAAGAGTGAACCGCAGAGATGATCTCAAGTGGAGTC 4317
Qy 5408 GCTTACCTTTGGCTTTGGCTTACAGAAACCTTATCTGTTTGTGACCCACTTCAACTCAC 5467
Db 4318 GCTTACCTTTGGCTTTGGCTTACAGAAACCTTATCTGTTTGTGACTCACTTCAACTCCC 4377
```

```
Qy 5468 TCGAAGTAATTGAGATCCAGGCACGCTCTCTCAGCAGGAGACCCCTGCCCCAGAGCTACTCG 5527
Db 4378 TGAAGTCAATTGAGATCCAGGCCAGATCTCTACTGGGAGCCCTGCCCCAGCATATCTCG 4437
Qy 5528 ACATCCCGAAACCCGCGCTACTGGGCCCTTGCATTTCTCTCAGGAGCGATTTACTTGGCGT 5587
Db 4438 AAATTCCAAACCTCTCTACTGGGCCCTTGCATTTCTCTCCGAGCGATTTACTTGGCGT 4497
Qy 5588 CCTCATACAGGATAAATTAAGGGTCAATTTGTGCAAGGAAACCTCTGTAAGGAGTCCG 5647
Db 4498 CCTCATACAGGACAAGTTAAGGGTCAATTTGTGCAAGGAAACCTCTGTAAGGAGTCA 4557
Qy 5648 GCACTGAACACCAACCCGGGGCCCTTCCACTTCCGCGAGCGCCCCCAACAGCGAGGCCAC 5707
Db 4558 GCACTGAGCAGCACCCGGTGTCTTCCACTTCCGCGAGCGCCCCCAACAGCGAGGCCAC 4617
Qy 5708 CCAGCTACACAGCAGCATCAACCAAGCGGTGCTTCCAGCCAGCGCCCGCCGAAAGGCC 5767
Db 4618 CAACATACAAACGAGCAGATCAACCAAGCGGTGCTTCCAGCCAGCGCCCGCCGAAAGGCC 4677
Qy 5768 CCAGCCACCCCGCAGAGCCAAAGCACACCCACCGCTACCAGAGCAGAGAGGGTCCGACAG 5821
Db 4678 CCAGCCACCCCGCAGAGCCAAAGCACACCGCTACCAGAGCAGAGAGGGTCCGACAG 4737
Qy 5822 AGCTGCGCAGGACAAGTCTCTTGGCGCCCTCTGGAGCGAGAGAGTCCCCCGCGCGA 5881
Db 4738 AGCTGCGCAGGACAAGTCTCTTGGCGCCCTCTGGAGCGGGAAGTCCCCAGGCCGA 4797
Qy 5882 TGCTCAGCAGCGGAGAGAGCGGTCCCCCGGAGGCTGTTTGAAGCAGCAGCAGCGGGGCC 5941
Db 4798 TGCTCAGCAGTGAAGAGAGCGGTCCCCAGGAGACTGTTTGAAGCAGCAGCAGCGGGGCC 4857
Qy 5942 GGCTGCTCGGGAGCGCTGAGGACCCCGCTGTCCAGGTGAACAAGGGGAAGAGGGCAGA 6001
Db 4858 GGCTGCTCGAGGAGCAGTGAAGACCCCACTGTCCCAGGTAAACAAGGTCTGGACACAGT 4917
Qy 6002 GTGCTCTCAAGTTT 6016
Db 4918 CTTAGTATAGTCT 4932
```

Search completed: July 3, 2004, 16:52:14  
Job time : 15441 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 12:14:14 ; Search time 9465 Seconds  
(without alignments)  
19431.714 Million cell updates/sec

Title: US-10-017-216-3  
Perfect score: 6159  
Sequence: 1 atgttgagttcaaatagg.....ttctgagacagattattgc 6159

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: em\_eetba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_estcl.\*

10: gb\_est2.\*

11: gb\_htc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rtd.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gssl.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	915.2	14.9	1011	13 B0070955	B0070955 AGENCOURT
2	815.6	13.2	956	13 B0071141	B0071141 AGENCOURT
3	810.2	13.2	830	13 B0181633	B0181633 AGENCOURT
4	750.6	12.2	849	12 B1253509	B1253509 602937370

5	716.4	11.6	920	13	BX342268	BX342268
6	692	11.2	879	13	BQ228524	BQ228524 AGENCOURT
7	688.8	11.2	958	11	BC031156	Mus muscu
8	629.2	10.2	652	10	BF905370	IL3-MT026
9	606.6	9.8	881	12	BG976452	602846269
10	589.8	9.6	746	14	CA749290	UI-M-FY0-
11	584.6	9.5	843	12	BG912161	BG912161 602812833
12	573.6	9.3	647	14	CF744580	UI-M-GV0-
13	572.8	9.3	730	14	CF723360	UI-M-GV0-
14	566.4	9.2	1085	12	BM904785	AGENCOURT
15	556.8	9.0	640	13	BQ807302	NISC_KK01
16	511.8	8.3	775	13	BUI05833	603005490
17	511	8.3	804	13	BUI236817	603411670
18	510.4	8.3	933	13	BUI232508	603408272
19	508.6	8.3	547	14	CF744826	UI-M-GV0-
20	491.6	8.0	875	13	BX723780	EX723780
21	490.8	8.0	875	13	BU900029	AGENCOURT
22	484.2	7.9	820	13	BQ941506	AGENCOURT
23	473	7.7	647	10	BB207065	BB207065
24	472.6	7.7	657	13	BQ831488	LI61n2149
25	471.6	7.6	681	10	CD348416	UI-M-FY0-
26	466.6	7.6	681	10	AM605350	QV3-DT004
27	451.4	7.3	459	13	BU429033	UI-HF-BN0
28	450.6	7.3	484	10	BE834307	RC1-OT008
29	445.6	7.2	891	12	BI558919	603240969
30	443.6	7.2	600	12	BG808196	2082-86 M
31	443.2	7.2	719	13	BUI339418	603515216
32	441.4	7.2	513	14	CF361392	827723 MA
33	437.8	7.1	860	14	CF223944	AGENCOURT
34	437.6	7.1	589	10	BF819093	QV2-CI010
35	435.2	7.1	1075	13	BQ425970	AGENCOURT
36	434.6	7.1	596	14	CF540358	UI-M-GV0-
37	431.2	7.0	497	10	BE909486	601502838
38	424.2	6.9	662	13	BUI04008	601502838
39	417.4	6.8	427	12	BG944539	axs5f111.x
40	417.2	6.8	749	13	BUI364866	603567832
41	415	6.7	678	13	BU403714	604137885
42	405.6	6.6	617	13	BQ417810	1k51b06.Y
43	398	6.5	936	12	BI758420	603026948
44	387.8	6.3	730	14	CA380902	660300 NC
45	384.2	6.2	1013	13	BUI08993	603110947

ALIGNMENTS

RESULT 1  
BQ070955  
LOCUS BQ070955 1011 bp mRNA linear EST 02-APR-2002  
DEFINITION AGENCOURT 685647 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:5923441  
5', mRNA sequence.

ACCESSION BQ070955 GI:19900001

VERSION BQ070955

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1011)

1 (bases 1 to 1011)

NH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.lnl.gov>

Plate: L1CM2091 row: p column: 02

High quality sequence stop: 634.

FEATURES		Location/Qualifiers	
QY	source	1. .1011	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5923441"	
QY		/tissue_type="neuroblastoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC 47"	
		/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."	
ORIGIN		Location/Qualifiers	
QY	Query Match	14.9%; Score 915.2; DB 13; Length 1011;	
	Best Local Similarity	97.9%; Pred. No. 4.6e-206;	
	Matches	946; Conservative 0; Mismatches 17; Indels 3; Gaps 2;	
QY	2653	CGGAGGTGAGTCTAGACACGAGGAGCAGAACTGGAGCTCAAGCGCAGCTCAGAG 2712	
	Db	1 CGGAGGTGAGTCTAGACACGAGGAGCAGAACTGGAGCTCAAGCGCAGCTCAGAG 60	
	QY	2713 CTACAGCTCTCCCTGAGGAGCGGAGTGCACAGTTCACAGCCCTGCAGGCTCGACGGCG 2772	
	Db	61 CTACAGCTCTCCCTGAGGAGCGGAGTGCACAGTTCACAGCCCTGCAGGCTCGACGGCG 120	
	QY	2773 GCCTGAGAGCGAGCTTCGCAGCGGGAAGACAGAGCTGGAAGACACACAGAGAAGCT 2832	
QY	Db	121 GCCTGAGAGCGAGCTTCGCAGCGGGAAGACAGAGCTGGAAGACACACAGAGAAGCT 180	
	QY	2833 GAAGAGGAGATCAGGACCTCAGGACATAGAGATGAATCCAGGCAATTTGATGCT 2892	
	Db	181 GAAGAGGAGATCAGGACCTCAGGACATAGAGATGAATCCAGGCAATTTGATGCT 240	
	QY	2893 CTTCTGAACAGCTGTACTTAATACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2952	
	Db	241 CTTCTGAACAGCTGTACTTAATACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 300	
QY	QY	2953 GACAGCGTGAATCAACACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 3012	
	Db	301 GACAGCGTGAATCAACACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 360	
	QY	3013 GGCAGCAAGCAGAGATTGTCAACTGCGAAGTGAAGTGAGCCATCTCCGCGGAGATC 3072	
	Db	361 GGCAGCAAGCAGAGATTGTCAACTGCGAAGTGAAGTGAGCCATCTCCGCGGAGATC 420	
	QY	3073 AGGAACGAGAGATGAGCTTACCGACGAGAGCAAAAGATGAGGCTCTGAGACCAAG 3132	
QY	Db	421 AGGAACGAGAGATGAGCTTACCGACGAGAGCAAAAGATGAGGCTCTGAGACCAAG 480	
	QY	3133 TGCACCATCTCTGAGGAAACAGCTCATGATTTGGAGGCGCTAAACGATGAGCTGTAGAA 3192	
	Db	481 TGCACCATCTCTGAGGAAACAGCTCATGATTTGGAGGCGCTAAACGATGAGCTGTAGAA 540	
	QY	3193 AAGAGCGCAGTGGAGCGCTTGGAGAGCGTCTCGGTGATGAGAAATCCAGTTTGAAG 3252	
	Db	541 AAGAGCGCAGTGGAGCGCTTGGAGAGCGTCTCGGTGATGAGAAATCCAGTTTGAAG 600	
QY	QY	3253 TGTGCGGTTGAGAGCTGAGAGAAATGCTGGACCGAGAAACAGACAGCGGCGAGAGCC 3312	
	Db	601 TGTGCGGTTGAGAGCTGAGAGAAATGCTGGACCGAGAAACAGACAGCGGCGAGAGCC 660	
	QY	3313 GATCAGCGGATCAGGAGTCTCCGAGGTGGAGCTGGAGCTGGAGTGAAGAGCAGCAAGGCT 3372	
	Db	661 GATCAGCGGATCAGGAGTCTCCGAGGTGGAGCTGGAGCTGGAGTGAAGAGCAGCAAGGCT 720	
	QY	3373 GAGATTCTCGTCTGACAGGCTCTCAAGAGCAGAGAGCTGAAGGCGGAGAGCTCTCTCT 3432	

Db	721	GAGATTCTCGTCTGACAGAGGCTCTCAAGAGCAGAAAGCTGGAGCGGAGAGCTCTCT 780	
QY	3433	GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCC- GAAGCTT 3491	
Db	781	GACAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAAGCTT 840	
QY	3492	ACAGCAGAAGCTGGAGACTGAACGAGAGCTCAACAGAGAGCTTCTGGAAGAGCAAGCCAA 3551	
Db	841	ACAGCAGAAGCTGGAGACTGGAGAGAGCTCAACAGAGAGCTCTGGANANCAANCCAA 900	
QY	3552	ATTACACAGAGATGACCTGGAGAGAGATCAATTTT--CCGTCTGATCAAGGACTG 3609	
Db	901	ATTACACAGAGATGACCTGGAGAGAGATCAATTTTTCGCGCTGGACTCAAGGACTG 960	
QY	3610	CAAGAA 3615	
Db	961	CCAGAA 966	
RESULT 2		Location/Qualifiers	
QY	BO071141	AGENCOURT 6853098 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5927502 5', mRNA sequence.	
	BO071141	BO071141.1 GI:19900187	
	VERSION	EST.	
	KEYWORDS	Homo sapiens (human)	
	SOURCE	Homo sapiens	
ORGANISM		Location/Qualifiers	
QY	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	AUTHORS	1. (bases 1 to 956)	
	TITLE	NIH-MGC http://mgi.nci.nih.gov/.	
	JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	
	COMMENT	Contact: Robert Strausberg, Ph.D. Email: ggaaps-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LNL) Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LCM2102 row: i column: 07 High quality sequence stop: 650.	
FEATURES		Location/Qualifiers	
QY	source	1. .956	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5927502"	
QY		/tissue_type="neuroblastoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC 47"	
		/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."	
ORIGIN		Location/Qualifiers	
QY	Query Match	13.2%; Score 815.6; DB 13; Length 956;	
	Best Local Similarity	97.1%; Pred. No. 2.2e-182;	
	Matches	874; Conservative 0; Mismatches 19; Indels 7; Gaps 4;	
QY	2653	CGGAGGTGAGTCTAGACACGAGGAGCAGAACTGGAGCTCAAGCGCAGCTCAGAG 2712	
	Db	1 CGGAGGTGAGTCTAGACACGAGGAGCAGAACTGGAGCTCAAGCGCAGCTCAGAG 60	
	QY	2713 CTACAGCTCTCCCTGAGGAGCGGAGTGCACAGTTCACAGCCCTGCAGGCTCGACGGCG 2772	
	Db	61 CTACAGCTCTCCCTGAGGAGCGGAGTGCACAGTTCACAGCCCTGCAGGCTCGACGGCG 120	
	QY	2773 GCCTGAGAGCGAGCTTCGCAGCGGGAAGACAGAGCTGGAAGACACACAGAGAAGCT 2832	
QY	Db	121 GCCTGAGAGCGAGCTTCGCAGCGGGAAGACAGAGCTGGAAGACACACAGAGAAGCT 180	
	QY	2833 GAAGAGGAGATCAGGACCTCAGGACATAGAGATGAATCCAGGCAATTTGATGCT 2892	
	Db	181 GAAGAGGAGATCAGGACCTCAGGACATAGAGATGAATCCAGGCAATTTGATGCT 240	
	QY	2893 CTTCTGAACAGCTGTACTTAATACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2952	
	Db	241 CTTCTGAACAGCTGTACTTAATACAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 300	
QY	QY	2953 GACAGCGTGAATCAACACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 3012	
	Db	301 GACAGCGTGAATCAACACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 360	
	QY	3013 GGCAGCAAGCAGAGATTGTCAACTGCGAAGTGAAGTGAGCCATCTCCGCGGAGATC 3072	
	Db	361 GGCAGCAAGCAGAGATTGTCAACTGCGAAGTGAAGTGAGCCATCTCCGCGGAGATC 420	
	QY	3073 AGGAACGAGAGATGAGCTTACCGACGAGAGCAAAAGATGAGGCTCTGAGACCAAG 3132	
QY	Db	421 AGGAACGAGAGATGAGCTTACCGACGAGAGCAAAAGATGAGGCTCTGAGACCAAG 480	
	QY	3133 TGCACCATCTCTGAGGAAACAGCTCATGATTTGGAGGCGCTAAACGATGAGCTGTAGAA 3192	
	Db	481 TGCACCATCTCTGAGGAAACAGCTCATGATTTGGAGGCGCTAAACGATGAGCTGTAGAA 540	
	QY	3193 AAGAGCGCAGTGGAGCGCTTGGAGAGCGTCTCGGTGATGAGAAATCCAGTTTGAAG 3252	
	Db	541 AAGAGCGCAGTGGAGCGCTTGGAGAGCGTCTCGGTGATGAGAAATCCAGTTTGAAG 600	
QY	QY	3253 TGTGCGGTTGAGAGCTGAGAGAAATGCTGGACCGAGAAACAGACAGCGGCGAGAGCC 3312	
	Db	601 TGTGCGGTTGAGAGCTGAGAGAAATGCTGGACCGAGAAACAGACAGCGGCGAGAGCC 660	
	QY	3313 GATCAGCGGATCAGGAGTCTCCGAGGTGGAGCTGGAGCTGGAGTGAAGAGCAGCAAGGCT 3372	
	Db	661 GATCAGCGGATCAGGAGTCTCCGAGGTGGAGCTGGAGCTGGAGTGAAGAGCAGCAAGGCT 720	
	QY	3373 GAGATTCTCGTCTGACAGGCTCTCAAGAGCAGAGAGCTGAAGGCGGAGAGCTCTCTCT 3432	

```
Db 1 CGGGAGGTCAGTCTAGACGACGAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCACAG 60
Qy 2713 CTACAGCTCTCCCTCAGGAGCGGAGTACAGTTGACAGCCCTGCAGGCTGCACGGCG 2772
Db 61 CTACAGCTCTCCCTCAGGAGCGGAGTACAGTTGACAGCCCTGCAGGCTGCACGGCG 120
Qy 2773 GCCCTGGAGGACCTTCGCCAGGCGAGACAGAGCTGGAGAGACACACAGCAGAGCT 2832
Db 121 GCCCTGGAGGACCTTCGCCAGGCGAGAGCAGAGCTGGAGAGACACACAGCAGAGCT 180
Qy 2833 GAAGAGGAGATCCAGGCACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2892
Db 181 GAAGAGGAGATCCAGGCACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 240
Qy 2893 CTTCTACAGCTGTAATCTAATCTACAGCTTCAGAGGAGAGCTTAACACAGCTCACCGAG 2952
Db 241 CTTCTACAGCTGTAATCTAATCTACAGCTTCAGAGGAGAGCTTAACACAGCTCACCGAG 300
Qy 2953 GACAAAGCTGAACTCAACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 3012
Db 301 GACAAAGCTGAACTCAACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 360
Qy 3013 GCGGCAACGAGCAGAGTGTACAACTGCGAAGTGAAGTGGACCAATCTCCGCGGAGATC 3072
Db 361 GCGGCAACGAGCAGAGTGTACAACTGCGAAGTGAAGTGGACCAATCTCCGCGGAGATC 420
Qy 3073 ACGGACGAGATCCAGCTTACCGCCAGAGCAAAAGATGGAGGCTCTGAAGACCAAG 3132
Db 421 ACGGACGAGATCCAGCTTACCGCCAGAGCAAAAGATGGAGGCTCTGAAGACCAAG 480
Qy 3133 TGCACCATGCTGGAGGAACAGTCTATGATTTGGAGGCGCCCTAAACGATGAGTGTGAGAA 3192
Db 481 TGCACCATGCTGGAGGAACAGTCTATGATTTGGAGGCGCCCTAAACGATGAGTGTGAGAA 540
Qy 3193 AAAGAGCGGAGTGGAGGCGCTGGAGGAGCGTCTCGGTGATGAGAAATCCAGTTTGAG 3252
Db 541 AAAGAGCGGAGTGGAGGCGCTGGAGGAGCGTCTCGGTGATGAGAAATCCAGTTTGAG 600
Qy 3253 TGTCCGGTTCGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 3312
Db 601 TGTCCGGTTCGAGAGCTGACAGAGATGCTGGACACCGAGAAACAGAGCGGCGAGAGCC 660
Qy 3313 GATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGCACAAGGCT 3372
Db 661 GATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGAGCACAAGGCT 720
Qy 3373 GAGATTCGCTCTCAGCAGGCTCTCAAGAGCAGAGAGCTGAAGCGGAGAGGCTCTCT 3432
Db 721 GAGATTCGCTCTCAGCAGGCTCTCAAGAGCAGAGAGCTGAAGCGGAGAGGCTCTCT 780
Qy 3433 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCGCGAAGC 3489
Db 781 GACAAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCGCGAAGC 840
Qy 3490 TTACAGCAGAA--GCTGGAGACTGAAC--GAGAGCTCAACAGAGGCTTCTGGAGAGCA 3545
Db 841 TTACAGCAGAAAGCTGGGAGACTGAACCGAAGACTCAACAGAGGCTTCTGGAGAGCA 900

RESULT 3
BUI81633
LOCUS BUI81633
DEFINITION AGENCOURT_7906225 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6140538
5', mRNA sequence.
ACCESSION BUI81633
VERSION BUI81633.1 GI:22695617
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 830)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
```

```
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13459 row: e column: 19
High quality sequence stop: 652.
FEATURES
Location/Qualifiers
1..830
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6140538"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo
Average insert size 1.75 kb. Library constructed by Life
Technologies."
```

## ORIGIN

```
Query Match 13.2%; Score 810.2; DB 13; Length 830;
Best Local Similarity 99.4%; Pred. No. 3.9e-181;
Matches 823; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 2387 TGGATTCCAAAGATCAGATCCCTGGAACAGAGGATTTGGAACTGTCTGAAGCCCAATAAAC 2446
Db 1 TGGATTCCAAAGATCAGATCCCTGGAACAGAGGATTTGGAACTGTCTGAAGCCCAATAAAC 60
Qy 2447 TTGAGCAAAATAGCAGTCTTTTACCCAAAGAAACATGAAGGCCCAAGAGAGATGATTT 2506
Db 61 TTGAGCAAAATAGCAGTCTTTTACCCAAAGAAACATGAAGGCCCAAGAGAGATGATTT 120
Qy 2507 CTGAACTCAGCAGCAAGAAATTTTACCTGGAGACAGAGCTGGAGCTTGGAGGCCCAGA 2566
Db 121 CTGAACTCAGCAGCAAGAAATTTTACCTGGAGACAGAGCTGGAGCTTGGAGGCCCAGA 180
Qy 2567 ACCGAAACTGGAGGAGCAGCTGGAGAGATCAGCCCAAGACCAAGCCACAGTGAACAAGATC 2626
Db 181 ACCGAAACTGGAGGAGCAGCTGGAGAGATCAGCCCAAGACCAAGCCACAGTGAACAAGATC 240
Qy 2627 GGCTGCTGGAACCTGGAGACAAAGATTGGGGAGGTTCAGTCTAGAGCAGGAGCAGAGAAC 2686
Db 241 GGCTGCTGGAACCTGGAGACAAAGATTGGGGAGGTTCAGTCTAGAGCAGGAGCAGAGAAC 300
Qy 2687 TGGAGCTCAAGCGCCAGCTCACAGAGTACAGCTCTCCCTGCGAGAGCGGAGTCAAGT 2746
Db 301 TGGAGCTCAAGCGCCAGCTCACAGAGTACAGCTCTCCCTGCGAGAGCGGAGTCAAGT 360
Qy 2747 TGACAGCCCTGCAAGCTGCAAGCGGCGCCCTGGAGAGCAGCTTTCGCGAGGAGAGCAG 2806
Db 361 TGACAGCCCTGCAAGCTGCAAGCGGCGCCCTGGAGAGCAGCTTTCGCGAGGAGAGCAG 420
Qy 2807 AGCTGGAGAGACCCACAGCAGAGAGCTGAAGAGAGATCCAGGCACTCACGCGCATAGAG 2866
Db 421 AGCTGGAGAGACCCACAGCAGAGAGCTGAAGAGAGATCCAGGCACTCACGCGCATAGAG 480
Qy 2867 ATGAAATCCAGCGCAAAATTTGATGCTTTTCGTAACAGCTGTACTGTAAATCACAGACTGG 2926
Db 481 ATGAAATCCAGCGCAAAATTTGATGCTTTTCGTAACAGCTGTACTGTAAATCACAGACTGG 540
Qy 2927 AGGAGCAGCTTAACAGCAGTCAAGCGAGCAACCGCTGAACTCAACACCAAAATTTCTACT 2986
Db 541 AGGAGCAGCTTAACAGCAGTCAAGCGAGCAACCGCTGAACTCAACACCAAAATTTCTACT 600
Qy 2987 TGTCCAAACCAACTCGATGAGGCTTCTGGCGCCCAACGACGAGATTTGTACAACTCGGAAGTG 3046
```

```

601 TGTCCTCAAACTCGATGAGGCTTCTGGGCGCCAAACGACGAGATTGTACAACTGCGAGTG 660
3047 AATGAGCACTATCCCGCGGAGATCAGGAAACGAGAGATGAGCTTACCAACCGAAGC 3106
661 AATGAGCACTATCCCGCGGAGATCAGGAAACGAGAGATGAGCTTACCAACCGAAGC 720
3107 AATGAGTGGAGGCTCTGAGACCACTGACCACTGCTGGAGGACAGGTCATGGATT-T-G 3165
721 AATGAGTGGAGGCTCTGAGACCACTGACCACTGCTGGAGGACAGGTCATGGATTGG 780
3166 GAGGCGCTTAAACGATGAGCTGTAGAAAAGAGCGGCGAGTGGAGGCC 3213
781 GAGGCGCTTAAACGATGAGCTGTAGAAAAGAGCGGCGAGTGGAGGCC 828

RESULT 4
BI253509 849 bp mRNA linear EST 17-JUL-2001
LOCUS 602973370F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112737 5',
DEFINITION mRNA sequence.
ACCESSION BI253509.1 GI:14805003
VERSION BI253509
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 849)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11273 row: 1 column: 18
High quality sequence start: 2
High quality sequence stop: 772.
Location/Qualifiers
1. 849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5112737"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 Kb. Library prepared by Life
Technologies."

FEATURES
source
Query Match 12.2%; Score 750.6; DB 12; Length 849;
Best Local Similarity 98.5%; Pred. No. 6.1e-167;
Matches 789; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

ORIGIN
Query 4659 ACTGCTTGAACCTCCTGCTGAACTGCAAGTGATGACCGTCTAGACATCAACTGCAC 4718
Db 3 ACCTGTCGGAACCTCCTGCTGAACTGGAAGTGATGACCGTCTAGACATCAACTGCAC 62
Query 4719 GCTGCCCTTCAGTGACAGAGTGGTGTGTGGGCGACCGAGGAGGGCTCTACGCCCTGAA 4778
Db 63 GCTGCCCTTCAGTGACAGAGTGGTGTGTGGGCGACCGAGGAGGGCTCTACGCCCTGAA 122
Query 4779 TGTCTTGAACAACTCCTTAACCATGTCTCCAGGAATTGAGCAGTCTTCCAAATTATAT 4838

```

```

Db 123 TGTCTTGAACAACTCCTTAACCATGTCTCCAGGAATCGGAGCAGTCTTCCAAATTTATAT 182
Query 4839 TATCAAGGACCTGAGAAAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGCTTGT 4898
Db 183 TATCAAGGACCTGAGAAAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTGCTTGT 242
Query 4899 GGAGCTGAAGAAAGTGAACAGTCCCTGGCGCCAGTCCCACCTGCTGCCAGCCCGACAT 4958
Db 243 GGAAGTGAAGAAAGTGAACAGTCCCTGGCGCCAGTCCCACCTGCTGCCAGCCCGACAT 302
Query 4959 CTCACCCAAATTTTGAAGCTGTCAAGGCTGCCACTTTGTTGGGCGAGGCAAGATTGA 5018
Db 303 CTCACCCAAATTTTGAAGCTGTCAAGGCTGCCACTTTGTTGGGCGAGGCAAGATTGA 362
Query 5019 GAACGGGCTCTGCATCTGTGCAGCCATGCCAGCAAGTCTGTCATTTCTCCGCTACAACGA 5078
Db 363 GAACGGGCTCTGCATCTGTGCAGCCATGCCAGCAAGTCTGTCATTTCTCCGCTACAACGA 422
Query 5079 AAACCTCAGCAAAATCTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTAT 5138
Db 423 AAACCTCAGCAAAATCTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTAT 482
Query 5139 CCACCTTCAACCAATACAGTATCTCTTGGAAACCAATTAATTTACGAAATCCAGATGAA 5198
Db 483 CCACCTTCAACCAATACAGTATCTCTTGGAAACCAATTAATTTACGAAATCCAGATGAA 542
Query 5199 GCAGTACACGCTCGAGGAATTCCTGGATAGAAATGACCATTCCTTGGCACCTCTGTGTT 5258
Db 543 GCAGTACACGCTCGAGGAATTCCTGGATAGAAATGACCATTCCTTGGCACCTCTGTGTT 602
Query 5259 TGCGGCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGCGAGCG 5318
Db 603 TGCGGCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGCAGGCGAGCG 662
Query 5319 AGAGGAGTACTTGTCTGTGTTTCCACGAATTTGGAGTGTCTCGTGGATTCTTACGGAAGACG 5378
Db 563 AGAGGAGTACTTGTCTGTGTTTCCACGAATTTGGAGTGTCTCGTGGATTCTTACGGAAGACG 722
Query 5379 T-AGCCGACAGACGATCTCAAGTGGAGTGCCTTACTTTGGCCTTTGCTTACAGAGAAC 5437
Db 723 TAAAGCCGACAGACGATCTCAAGTGGAGTGCCTTACTTTGG-CTTTGCCCTACCGAGAA- 780
Query 5438 CCTATCTGTTGTGACCCACT 5458
Db 781 CCTATCTGTTGGACCCATT 801

RESULT 5
BX342268 920 bp mRNA linear EST 02-MAY-2003
LOCUS BX342268
DEFINITION CDNA clone CSDBK011YB20 5-PRIME, mRNA sequence.
ACCESSION BX342268.1 GI:30334095
VERSION BX342268
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 920)
AUTHORS Li,W.B., Gruber,C., Jesse,J., and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6533.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSDBK011YB20
Feng Liang Email : fliang@lifetech.com URL :

```

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DK011DA10QPL

FEATURES

source  
1..920  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK011YB20"  
/cell\_type="HELA CELLS COT 25-NORMALIZED"  
/cell\_lines="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 11.6%; Score 716.4; DB 13; Length 920;  
Best Local Similarity 95.9%; Pred. No. 9e-159;  
Matches 746; Conservative 0; Mismatches 31; Indels 1; Gaps 1;  
QY 5265 CTCCTTCCAAACAGCTCCCTGCTCAATCGTGCAGGTGAACAGCGAGCGAGGAGAGGA 5324  
DB 66 CTCCTTCCAAACAGCTCCCTGCTCAATCGTGCAGGTGAACAGCGAGCGAGGAGAGGA 125  
QY 5325 GTACTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGCCG 5384  
DB 126 GTACTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGCCG 185  
QY 5385 CACAGAGGATCTCAAGTGAGTGCCTTACCTTTGGCTTGTGCTACAGAGACCTTACT 5444  
DB 186 CACAGAGGATCTCAAGTGAGTGCCTTACCTTTGGCTTGTGCTACAGAGACCTTACT 245  
QY 5445 GTTGTGACCCACTTCAACTCACTCAAGTAAATTTAGATTCAGGACACGCTCTCTCAGCAGG 5504  
DB 246 GTTGTGACCCACTTCAACTCACTCAAGTAAATTTAGATTCAGGACACGCTCTCTCAGCAGG 305  
QY 5505 GACCCCTGCCAGGCTACTTGACATCCGAAACCGGGCTACTGCGGCCCTGCCATTC 5564  
DB 306 GACCCCTGCCAGGCTACTTGACATCCGAAACCGGGCTACTGCGGCCCTGCCATTC 365  
QY 5565 CTCAGAGGAGTATTACTTGGCGTCTCTATACCAGGATAAATTAAGGGTCATTGTGCGAA 5624  
DB 366 CTCAGAGGAGTATTACTTGGCGTCTCTATACCAGGATAAATTAAGGGTCATTGTGCGAA 425  
QY 5625 GGGAACTCTGTGAAGAGTCCGGCACTGAAACACCGGGGCCGCTCAGCTCCCGCAG 5684  
DB 426 GGGAACTCTGTGAAGAGTCCGGCACTGAAACACCGGGGCCGCTCAGCTCCCGCAG 485  
QY 5685 CAGCCCAACAGCGAGGCCACCCAGCTACACAGGACACATCACCAAGCGGTGCGCTC 5744  
DB 486 CAGCCCAACAGCGAGGCCACCCAGCTACACAGGACACATCACCAAGCGGTGCGCTC 545  
QY 5745 CAGCCCAAGCGCGGAGGCCACCCAGCTACCCAGGAGCCAGCACACCCACCGCTA 5804  
DB 546 CAGCCCAAGCGCGGAGGCCACCCAGCTACCCAGGAGCCAGCACACCCACCGCTA 604  
QY 5805 CCGGAGGGGGCGGACCGAGCTCGGAGGACAAAGTCTCTGCGCGCCCTTGAGCGGAGA 5864  
DB 605 CCGGAGGGGGCGGACCGAGCTCGGAGGACAAAGTCTCTGCGCGCCCTTGAGCGGAGA 664  
QY 5865 GAAATCCCCCGCGGATGCTCAGACCGGAGAGAGCGGTCCCCCGGAGGCTGTTGA 5924  
DB 665 GAAATCCCCCGCGGATGCTCAGACCGGAGAGAGCGGTCCCCCGGAGGCTGTTGA 724  
QY 5925 AGACAGCAGAGGGGGCGGCTGCTGCGGGAGCGGTGAGGACCGCGTGTCCAGGTGAA 5984  
DB 725 AGACAGCAGAGGGGGCGGCTGCTGCGGGAGCGGTGAGGACCGCGTGTCCAGGTGAA 784  
QY 5985 CAGGGAAGGGGCGAGGTCCTCTCAAGTTTTCAGGTTTACACTGTTCACCTATTAT 6042  
DB 785 CAAGGTCTGGGACCACTCTTCAGTATATAATCTCAGCCAGAAAAACCAACTCTCTACT 842

RESULT 6

LOCUS BQ228524  
DEFINITION AGENCOURT 7600932 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6061485  
5', mRNA sequence.  
ACCESSION BQ228524  
VERSION BQ228524.1 GI:20409924  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 879)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-f@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM1331 row: g column: 22  
High quality sequence stop: 636.  
Location/Qualifiers

FEATURES

source  
1..879  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6061485"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 Kb. Library constructed by Life  
Technologies."

ORIGIN

Query Match 11.2%; Score 692; DB 13; Length 879;  
Best Local Similarity 100.0%; Pred. No. 5.6e-153;  
Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3946 CGGAGGAAAGCTGCCACCGCAAGCAACGAGCAACCCATCCAGCCAGCCAGCCAGCC 4005  
DB 1 CGGAGGAAAGCTGCCACCGCAAGCAACGAGCAACCCATCCAGCCAGCCAGCCAGCC 60  
QY 4006 CGGAGGAGAGATCGGCATGTCGGCATCGTGGGTGCGGAGCAGCAGCCAGCCAGCC 4065  
DB 61 CGGAGGAGAGATCGGCATGTCGGCATCGTGGGTGCGGAGCAGCAGCCAGCCAGCC 120  
QY 4066 ATGAGCTGTGCGCCCGCCATCCAGCGCAGAAAGAGTCTTCACTCCAGAGGAATTT 4125  
DB 121 ATGAGCTGTGCGCCCGCCATCCAGCGCAGAAAGAGTCTTCACTCCAGAGGAATTT 180  
QY 4126 AGTCGGCGTCTTAAGGAACGCGATGCACCAATATTCCTACCGATTCAACGTAGGACTG 4185  
DB 181 AGTCGGCGTCTTAAGGAACGCGATGCACCAATATTCCTACCGATTCAACGTAGGACTG 240  
QY 4186 AACATGCGAGCCCAAAAGT 4245  
DB 241 AACATGCGAGCCCAAAAGT 300  
QY 4246 TCCAAATGTCTCGAATGTCAAGGTGATGTGACACCCCAAGTGTCCAGTGTCTTGCAGCC 4305  
DB 301 TCCAAATGTCTCGAATGTCAAGGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360  
QY 4306 ACCTGCGGCTGCTGT 4365

Db 361 ACCTGGGGCTTGCTGCTGTAATATGCCACACACTTCACCGAGGCTTCTGCCGTGACAAA 420  
 Qy 4366 ATGAATCTCCCGAGTCTCCAGACCAAGGAGCCGACGAGCTTGCACTGGAAGGGTGG 4425  
 Db 421 ATGAATCTCCCGAGTCTCCAGACCAAGGAGCCGACGAGCTTGCACTGGAAGGGTGG 480  
 Qy 4426 ATGAAGTGGCCAGGATTAACAACAGGAGACACCAAGGCTGGACAGGAGTACATTGTC 4485  
 Db 481 ATGAAGTGGCCAGGATTAACAACAGGAGACACCAAGGCTGGACAGGAGTACATTGTC 540  
 Qy 4486 CTGAGGAGTCAAAAGTCTCATTTATGACAATGAAGCCAGACGAAGCTGCACAGAGCCG 4545  
 Db 541 CTGAGGAGTCAAAAGTCTCATTTATGACAATGAAGCCAGACGAAGCTGCACAGAGCCG 600  
 Qy 4546 GTGGAAGAAATTGAGCTGTCCTTCCGACGGGATGTATCTATTCAATGTCCTGTTGTT 4605  
 Db 601 GTGGAAGAAATTGAGCTGTCCTTCCGACGGGATGTATCTATTCAATGTCCTGTTGTT 660  
 Qy 4606 GCTTCCGAATCTCGCAATATACAGCCAAAGCAGA 4637  
 Db 661 GCTTCCGAATCTCGCAATATACAGCCAAAGCAGA 692

RESULT 7  
 BC031156  
 LOCUS 958 bp mRNA linear HTC 04-MAR-2003  
 DEFINITION Mus musculus, Similar to citron, clone IMAGE:4976752, mRNA.  
 ACCESSION BC031156  
 VERSION BC031156.1 GI:21411076  
 KEYWORDS HTC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 958)  
 DIRECT SUBMISSION  
 SUBMITTED (03-JUN-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 CONTACT: MGC help desk  
 EMAIL: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
 TISSUE: Tissue Procurement: Jeffrey Green M.D.  
 CDNA LIBRARY PREPARATION: Life Technologies, Inc.  
 CDNA LIBRARY ARRANGED BY: The I.M.A.G.E. Consortium (LLNL)  
 DNA SEQUENCING BY: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 WEB SITE: <http://www.shgc.stanford.edu>  
 CONTACT: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 DICKSON, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAP Plate: 59 Row: J Column: 6  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis  
 This clone has the following problem: retained intron.

FEATURES  
 source  
 Location/Qualifiers  
 1..958  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4976752"  
 /tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating  
 ductal carcinoma. 5 month old virgin mouse."  
 /clone\_lib="NCI CGAP\_Mam6"  
 /lab\_host="DH103"  
 /note="Vector: pCMV-SPORT6"

ORIGIN  
 Query Match 11.2%; Score 688.8; DB 11; Length 958;  
 Best Local Similarity 85.4%; Pred. No. 3.5e-152;  
 Matches 781; Conservative 0; Mismatches 127; Indels 6; Gaps 1;  
 Qy 5149 AATTACAGTATCTTCATTGGAACCAATAAATTTACGAAATCGACATGAAGCAGTACACG 5208  
 Db 40 AAGAAACGCAATCTTCATTGGGACCAACAAATTCATGAGATCGACATGAAGCAGTACACG 99  
 Qy 5209 CTCGAGGAATTCCTTGGATTAAGAATGACATTCCTTGGCACCTGCTGTGTTTGGCGCTCT 5268  
 Db 100 CTTGATGATTCCTGGACAGACGACCATTCCTTGGCACTGCTGTGTTTGGCGCTCTCTCG 159  
 Qy 5269 TCCAAACAGCTTCCCTGCTCAATCGTCAGTGAAACAGCCAGGAGCAGCAGAGAGATAC 5328  
 Db 160 TCCAAACAGCTTCCCTGCTCAATTCGAGCGCAACAGCGCGGAGAGAGATAC 219  
 Qy 5329 TTGCTGTGTTTCCACGAAATTTGGAGTCTTCGTGATTTCTTACGAAAGACGTAGCCGACA 5388  
 Db 220 CTGCTGTGTTTCCACGAAATTTGGGTGTTCTGATTTCTTACGAAAGCAGTACCGCACA 279  
 Qy 5389 GACGATCTCAAGTGGAGTCGCTTACCTTTGGCCTTTTGCCTTACAGAAACCTATCTGTTT 5448  
 Db 280 GATGATCTTAAGTGGAGTCGCTTACCTTGGCCTTGGCCTTACAGAGAACCTTATCTGTTT 339  
 Qy 5449 GTGACCCACTTCACTCACTCGAAGTAATTGAGATCCAGGACGCTCCTCAGCAGGAC 5508  
 Db 340 GTGACTCACCTTCACTCCTCTGGAAGTCATTGAGATCCAGGCCAGATCCTCCTCTGGGAGC 399  
 Qy 5509 CCTGCCGAGCGTACCTGGACATCCCGAACCCCGGCTACCTGGCCCTTGCCATTTCTCTCA 5568  
 Db 400 CTTGCCCGAGCATATCTGGAATTCGAAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 459  
 Qy 5569 GGACCGATTTACTTGGCGTCTCTATACCAGGATAAATTAAGGTCTATTGCTCAGAGGA 5628  
 Db 460 GGACCGATTTACTTGGCGTCTCTATACCAGGATAAATTAAGGTCTATTGCTCAGAGGA 519  
 Qy 5629 AACTCTGTGAAGGAGTCCGGCACTGAAACACACCGGGGCCGCTCCACCTCCCGCAGCAGC 5688  
 Db 520 AACTCTGTGAAGGAGTCCGGCACTGAGCAGCACCGGGTGCCTCCACCTCCCGCAGCAGC 579  
 Qy 5689 CCAACAGGAGGCCCAACCCAGTACACGAGACATACCAAGCGCTGGCTCCAGC 5748  
 Db 580 CCAACAGGAGGCCCAACCAACATACACGAGCACATCAACCAAGCGCTGGCTCCAGC 639  
 Qy 5749 CCAGCGCGCCCGAAGCCCGCCAGCCAGAGCCAGAGCCAGAGCACCCACCGCTACC-- 5806  
 Db 640 CCGCGCGCACCGAAGGCCCGCCAGCCAGCCAGAGCCAGAGCACCGACCCGCTACCGA 699  
 Qy 5807 ----GCGAGGGCGGACCCGAGCTGCGCAGGACAAGTCTCTGCGCGCGCCCTCTGGAGCGA 5862  
 Db 700 GACAGAGAGGGTCCGACAGAGCTGCGCAGGACAAGTCTCCAGGCGCGCTCTGGAGCGG 759  
 Qy 5863 GAGAGTCTCCCGCGGATGCTCAGCACCGGAGAGCGGCTCCCGGAGGCTGTTT 5922  
 Db 760 GAGAAGTCTCCCGCGGATGCTCAGCACTAGGAGAGCGGCTCCCGGAGGAGTCTT 819  
 Qy 5923 GAAGACAGCAGCAGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCCGAGGTG 5982  
 Db 820 GAAGACAGCAGCAGGCGCGCTGCTGCGAGGAGCAGTGGAGCCCGCTGTCTCCAGTT 879  
 Qy 5983 AACAGGGAAGGCGGAGAGTCTCTCAAGTTTTCAGGTTTACGTTTACACTGTCCACTATTAT 6042  
 Db 880 AACAGGTCCTGGAGCCAGTCTTCAGTATAGTCTAGCCCAAAAAA 939  
 Qy 6043 GACTGGAATAAAAA 6056  
 Db 940 AAAAAA 953

RESULT 8  
 BF905370/c

LOCUS BF905370 652 bp mRNA linear EST 18-JAN-2001  
DEFINITION IL3-MT0267-261200-410-H07 MT0267 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF905370  
VERSION BF905370.1 GI:12296829  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 652)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PURVED 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=IL3&t2=IL3-MT0267-  
261200-410-H07&t3=2000-12-26&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 624.  
Location/Qualifiers  
1..652  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="MT0267"  
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."  
ORIGIN  
Query Match 10.2%; Score 629.2; DB 10; Length 652;  
Best Local Similarity 98.0%; Pred. No. 4e-138;  
Matches 637; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 2477 GGAACATGAAGCCCAAGAGAGATGATTTCTGAACCTCGGCAACAGAAATTTACCTGG 2536  
DB 652 GGAACATGAAGCCCAAGAGAGATGATTTCTGAACCTCGGCAACAGAAATTTACCTGG 593  
QY 2537 AGACACAGCTGGGAATTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTGGAGAAGA 2596  
DB 592 AGACACAGCTGGGAATTTGGAGGCCAGAACCGAAACTGGAGGAGCAGCTGGAGAAGA 533  
QY 2597 TCAGCCACCAAGACCAAGTGCAGAAATCGGCTGCTGGAACTGGAGCAAGATTTCGGGG 2656  
DB 532 TCAGCCACCAAGACCAAGTGCAGAAATCGGCTGCTGGAACTGGAGCAAGATTTCGGGG 473  
QY 2657 AGTCACTAGACGACGAGGAGGAGAACTGGAGCTCAAGCGCCAGCTCAGAGAGCTAC 2716  
DB 472 AGTCACTAGACGACGAGGAGGAGAACTGGAGCTCAAGCGCCAGCTCAGAGAGCTAC 413  
QY 2717 AGCTCTCCCTGCAGGAGCGGAGTACAGTTCAGAGCCCTGCAGGCTGCACGGCGGCC 2776

DB 412 AGCTCTCCCTGCAGGAGCGGAGTACAGTTGACAGCCCTGCAGGCTGCACGGCGGCC 353  
QY 2777 TGGAGAGCCAGCTTGGCAGGCGGAGACAGAGCTGGAAGAGACACAGAGAGAGCTGAAG 2836  
DB 352 TGGAGAGCCAGCTTGGCAGGCGGAGACAGAGCTGGAAGAGACACAGAGAGAGCTGAAG 293  
QY 2837 AGGAGATCCAGGCACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCTTC 2896  
DB 292 AGGAGATCCAGGCACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCTTC 233  
QY 2897 GTAACAGCTTACTGTAAATCACAGACCTGGAGAGAGCTTAAACCAAGCTGACGAGGACA 2956  
DB 232 GTAACAGCTTACTGTAAATCACAGACCTGGAGAGAGCTTAAACCAAGCTGACGAGGACA 173  
QY 2957 ACCTGGAACTCACACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCG 3016  
DB 172 ACCTGGAACTCACACACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCG 113  
QY 3017 CCAACGACGAGATGTACAACTCGAAGTGAAGTGAACCATCTCCGCGGAGAGTACACGG 3076  
DB 112 CCAACGACGAGATGTACAACTCGAAGTGAAGTGAACCATCTCCGCGGAGAGTACACGG 53  
QY 3077 AACGAGATGCGAGCTTACCAGCCAGAGCAACGATGGAGGCTCTGAAG 3126  
DB 52 AACGAGATGCGAGCTTACCAGCCAGAGCAACGATGGAGGCTCTGAAG 3  
RESULT 9  
LOCUS BG976452 881 bp mRNA linear EST 12-JUN-2001  
DEFINITION 602846269F1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:4976752 5',  
mRNA sequence.  
ACCESSION BG976452  
VERSION BG976452.1 GI:14364089  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 881)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbe-x@mail.nih.gov  
Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM10970 row: j column: 17  
High quality sequence stop: 841.  
Location/Qualifiers  
1..881  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4976752"  
/sex="female, virgin"  
/tissue\_type="infiltrating ductal carcinoma"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Mam6"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Jeffrey Green, M.D., NIH"  
ORIGIN  
Query Match 9.8%; Score 606.6; DB 12; Length 881;



```
Db 301 CAGGTGGTGGAGTTGGCGGTGAAGGAACACAGGCCGCGAGATTCTGTCTCGACAGGCT 360
QY 3397 CTCRAAGAGCAGAAAGCTCAAGGCCGAGAGCCTCTCTGACAAAGCTCAATGAACTCGAGAAG 3456
Db 361 CTCRAAGAGCAGAAAGCTCAAGGCCGAGAGCCTGTGCGACAAAGCTCAACGACCTCGAGAAG 420
QY 3457 AAGCATGCTATGCTTGAATGAATGCCGAGACTTACAGCAGAAAGCTGAGACACTGGAACGA 3516
Db 421 AAGCAGCGCATGCTGGAGATGAACGCCCGGAGCTTACAGCAGAAACTAGACACAGAGCGG 480
QY 3517 GAGCTCAACACAGAGCTTCTGGAGCAGCAAGCAAAATTACAGCAGCAGATGGACCTGCGAG 3576
Db 481 GAGCTCAACACAGAGCTTCTGGAGCAGCAAGCAAAATTGAGCAGCAGATGGACCTGCGAG 540
QY 3577 AAAATACATTTTCGCTGCTGACTCAAGGACTGCAAGAAAGCTCTAGATCGGGCTGATCTA 3636
Db 541 AAGAACCAATCTTTCAGACTGACGAAAGGCTGAGAGGGCTGAGAGGGCTGAGAGGGCGGATCTG 600
QY 3637 CTGAGACAGAAAGAGTCTTGGAGTATCAGCTGGGAACATTCAGGTTCTCTATTCT 3696
Db 601 CTGAAAGCAGAAAGAGGAGCCTTGAGTACCAGCTGGGAAACATTCAGGTTCTCTATTCT 660
QY 3697 CATGAAAAGGTGAAATGGAAGGCACTATTTCTCAACAAACCAAACTCAATGATTTCTG 3756
Db 661 CACGAGAAAGTGAAATGGAAGGCACTATCTCTCAGCANA-CANACTCATTTGATTTCTG 719
QY 3757 CAGGCCAAATGGACCAACCTGCTTAAAA 3784
Db 720 CAAG-CANAATGGACCAACCTGCTTAAAA 746

RESULT 11
LOCUS BG912161
DEFINITION 602812B33F1 NCI_CGAP Brn67 Homo sapiens cDNA clone IMAGE:4944657
5', mRNA sequence.
ACCESSION BG912161
VERSION BG912161.1 GI:14292637
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10890 row: a column: 10
High quality sequence stop: 778.
Location/Qualifiers
1..843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4944657"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="pHL10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Brn67"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source
```

```
Query Match 9.5%; Score 584.6; DB 12; Length 843;
Best Local Similarity 92.8%; Pred. No. 2e-127;
Matches 669; Conservative 0; Mismatches 44; Indels 8; Gaps 5;

QY 4636 GAAAAAGCAGAGCTCATGCTAAACTGCTTGGAAACTCCCTGCTGAACTGGAAGTGAT 4695
Db 122 GAAAAAGCAGAGCTCATGCTAAACTGCTTGGAAACTCCCTGCTGAACTGGAAGTGAT 181
QY 4696 GACCGTCTAGACATGAATGACGCTGCCCTTCAAGTACAGAGTGGTGTGTGGGCAAC 4755
Db 182 GACCGTCTAGACATGAATGACGCTGCCCTTCAAGTACAGAGTGGTGTGTGGGCAAC 241
QY 4756 GAGGAAGGCTCTACGCCCTGAATGCTTGAATACTCCCTAACCATGCTCCCGAGAAAT 4815
Db 242 GAGGAAGGCTCTACGCCCTGAATGCTTGAATACTCCCTAACCATGCTCCCGAGAAAT 301
QY 4816 GGAGCAGCTCTTCCAAATTTATATTATCAAGGACCTCGAGAAGCTACTCATATAGCAGGA 4875
Db 302 GGAGCAGCTCTTCCAAATTTATATTATCAAGGACCTCGAGAAGCTACTCATATAGCAGGA 361
QY 4876 GAAGAGCGGCACTGTGCTTGTGGAAGTGAAGAAAGTGAACAGTCCCTGGCCCAAGTCC 4935
Db 362 GAAGAGCGGCACTGTGCTTGTGGAAGTGAAGAAAGTGAACAGTCCCTGGCCCAAGTCC 421
QY 4936 CACCTGCTGCCAGCCGACATCTCACCAACATTTTGAAGCTCTCAAGGGCTGCCAC 4995
Db 422 CACCTGCTGCCAGCCGACATCTCACCAACATTTTGAAGCTCTCAAGGGCTGCCAC 481
QY 4996 TTGTTTGGGSCAGGCAAGATTGAGAAAGGCTCTGCATCTGTGACGCAATGCCAGCAAA 5055
Db 482 TTGTTTGGGSCAGGCAAGATTGAGAAAGGCTCTGCATCTGTGACGCAATGCCAGCAAA 541
QY 5056 GTGCTCATTTCTCCGTACACGAAACCTCAGCAATATCTGCATCCGGAAGAGATAGAG 5115
Db 542 GTGCTCATTTCTCCGTACACGAAACCTCAGCAATATCTGCATCCGGAAGAGATAGAG 601
QY 5116 ACCTCAGAGCCCTGCAGCTGTATCCACTTCCACCAATTACAGTATCTCTCATTT-GGAACCAA 5174
Db 602 ACCTCAGAGCCCTGCAGCTGTATCCACTTCCACCAATTACAGTATTTCTCATTTGGGAACCAA 661
QY 5175 TAAATCTAGC-AAATCGACATGAA-GCAGTACACCTCGAGGAAT--CTTGATAAGA 5230
Db 662 TAAATCTAGCAAAATCGACATGAGGCACTACACGCTCGAGGAATTTCTTGATAACGA 721
QY 5231 ATGACCATTTCTTGGCACCTGCTGTG---TTTCCGCCCTCTTCCAAACAGCTTCCCTGTCT 5287
Db 722 ATGACCATTTCTTGGCACCTGCTGTG---TTTCCGCCCTCTTCCAAACAGCTTCTGTCTC 781
QY 5288 CAATCTGCAAGTGAACAGCGCAGGCGCAGAGAGAGTACTTGTGTGTTTCCAGCAAT 5347
Db 782 AATCGTGGCAGGTGAACAGGCGCAGGCGCAGAGAGAGTATTGCTGTGTTCCCGCACT 841
QY 5348 T 5348
Db 842 T 842

RESULT 12
LOCUS CF744580
DEFINITION UI-M-GVO-clt-n-23-0-UI.r1 NIH_BMAP_GVO Mus musculus cDNA clone
IMAGE:30617710 5', mRNA sequence.
ACCESSION CF744580
VERSION CF744580.1 GI:37640920
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mousefl.html  
(This clone was contributed by the Brain Molecular Anatomy Project (BMAP))

Seq primer: pYX-5.

## FEATURES

source  
Location/Qualifiers  
1. .647  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30617710"  
/tissue\_type="whole brain"  
/dev\_stage="1.5, and 15 days newborn"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_GV0"  
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: Ecor I;  
Site 2: Not I; The library was constructed according  
Bonaïdo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated with  
Ecor I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CGAATCGAAT. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

## ORIGIN

Query Match 9.3%; Score 573.6; DB 14; Length 647;  
Best Local Similarity 93.8%; Pred. No. 6.9e-125;  
Matches 608; Conservative 0; Mismatches 39; Indels 1; Gaps 1;  
QY 1131 CACCCTCAAGTCGACGATGACACCTCCAAATTTTGTATGACGAGAGAAATTCGTGGGT 1190  
DB 1 CACCCTCAAGTCGACGATGACACCTCCAAATTTTGTATGACGAGAGAAATTCGTGGGT 60  
QY 1191 TTTCATCTCTCGTGCCAGTCGAGCCCTCAGGCTTCTCGGTGGAAGAACTGCCGTTTGT 1250  
DB 61 TTTCATCTCTGTGTGCCAGTCGAGCCCTCGGCTTCTCAGCGAAGAGCTGCCGTTTGT 120  
QY 1251 GGGGTTTTCTGACAGCAAGGCACTGGGATTTCTTGGTAGATCTGAGTCTGTGTGTCGGG 1310  
DB 121 GGGATTTTCTGACAGCAAGGCACTGGGATTTCTTGGTAGATCTGAGTCTGTGTGTCGAG 180  
QY 1311 TCTGGACTCCCTGCCAGCACTAGCTCCATGGAAGAAACTTCTCATCAAAAGCAAGA 1370  
DB 181 TCTGGACTCCCTGCCAGGTTAGCTTCATGGAAGAAACTTCTCATCAAAAGCAAGA 240  
QY 1371 GCTACAAGACTCTCAGGCAAGTGTCAAGAATGGAGCAGGAAATGACCCGGTTACATCG 1430  
DB 241 GCTCCAAGACTCCAGGCAAGTGTCAAGAATGGAGCAGGAAATGACCCGGTTACATCG 300  
QY 1431 GAGAGTGTGAGAGGTGGAGCTGCTTAGTCAGAGAGGTGGAGCTGAGGCTCTGA 1490  
DB 301 CAGAGTGTGAGAGGTGGAGCTGCTTAGTCAGAGAGGTGGAGCTGAGGCTCTGA 360  
QY 1491 GACTCAGAGATCCCTCTGGAGCAGGACCTTCTACCTACATCAAGAAATGCAAGTGT 1550  
DB 361 GACTCAGAGATCCCTCTGGAGCAGGACCTTCTACCTACATCAAGAAATGCAAGTGT 420  
QY 1551 AAAGCGAAGTTTGGAGCAAGCAGGATGGAGGTGTCCCGAGGAGTGAACAAGCACTGCA 1610

Db 421 AAAGCGAAGTTTGGAGCAAGCGGATGGAGGTGTCCCGAGGATGACAAAGCTCTGCA 480  
QY 1611 GCTTCTCCATGATATCAGAGCAGAGCCGAGAGCTCCAAAGAAATCAAGAGCAGAGTA 1670  
DB 481 GCTTCTCCAGACATCCGAGCAGAGCCGAGAGCTCCAGGAGATCAAGGAGCAGGTA 540  
QY 1671 CCAGGCTCAAGTGAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTC 1730  
DB 541 CCAGGCTCAGTGGAGAGATGAGGCTGATGATGATCAGTGGAGAGA-CTGTGTC 599  
QY 1731 AGCAAGAGCAGGAGTGTCTTACGAATCTGAGTGGAGAGATCTCG 1778  
DB 600 AGCCCGCAGACGAGCGATCTTACGAGTCTGAGCTGAGGAGTCTCG 647

## RESULT 13

CF723360 730 bp mRNA linear EST 09-OCT-2003  
LOCUS UI-M-GV0-cjh-j-20-0-UI.r1 NIH\_BMAP\_GV0 Mus musculus cDNA clone  
DEFINITION IMAGE:30546187 5', mRNA sequence.  
ACCESSION CF723360  
VERSION CF723360.1 GI:37597528  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 730)  
NIH-MGC http://mgs.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mousefl.html  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

## FEATURES

source  
Location/Qualifiers  
1. .730  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30546187"  
/tissue\_type="whole brain"  
/dev\_stage="1.5, and 15 days newborn"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_GV0"  
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: Ecor I;  
Site 2: Not I; The library was constructed according  
Bonaïdo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated with  
Ecor I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CGAATCGAAT. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

## ORIGIN

Query Match 9.3%; Score 572.8; DB 14; Length 730;  
Best Local Similarity 88.7%; Pred. No. 1.2e-124;

Matches 643; Conservative 0; Mismatches 75; Indels 7; Gaps 2;									
QY	5194	ATGAAGCAGTACAGCTCGAGAAATTCCTGATAGAATGACCAATTCCTTGGACCTGCT	5253						
Db	1	ATGAAGCAGTACAGCTTGACGATTCCTGACAAAGAACACCAATTCCTTGGCACCCTGCT	60						
QY	5254	GTGTTTCGGCTCTTCCACAGCTTCCTGCTCTCAATCGTGCAGGTGAACAGCGCAGGG	5313						
Db	61	GTGTTTCGGCTCTTCCACAGCTTCCTGCTCTCAATCGTGCAGGTGAACAGCGCAGGG	120						
QY	5314	CAGCGAGAGGAGTACTTCTGTTTCCACGAAATTTGGAGTGTTCGTGGATCTTACCGA	5373						
Db	121	CAGCGAGAGGAGTACTTCTGTTTCCACGAAATTTGGAGTGTTCGTGGATCTTACCGA	180						
QY	5374	AGACGTAGCCGACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCTTTGCCCTACAGA	5433						
Db	181	AGACGTAGCCGACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCTTTGCCCTACAGA	240						
QY	5434	GAACCTATCTGTTTGTGACCCACTTCAACCTCACTCGAAGTAAATTGAGATCCAGGCACGC	5493						
Db	241	GAACCTATCTGTTTGTGACCTCACTTCAACTCCCTGAGTCAATTGAGATCCAGGCACGC	300						
QY	5494	TCCTCAGCAGGAGCCCTGCGCAGCGTACCTGAGATCCCGAACCCCGCGTACTCGGC	5553						
Db	301	TCCTCAGCAGGAGCCCTGCGCAGCGTACCTGAGATCCCGAACCCCGCGTACTCGGC	360						
QY	5554	CCTGCCATTTCTCAGGAGCGATTTACTTGGCGTCTCATACCAAGGATAAAATTAAGGTC	5613						
Db	361	CCCGGATTTCTCCGAGCGATTTACTTGGCGTCTCATACCAAGGATAAAATTAAGGTC	420						
QY	5614	ATTTCTCAAGGGAACCTCGTGAAGGAGTCGCGCACTGAACACCCAGCGGGCCGCTCC	5673						
Db	421	ATATCTGCAAGGAAACCTCGTGAAGGAGTCAGGCACTGAGCAGCACCGGTGCCCTCC	480						
QY	5674	ACCTCCGCGAGAGCCCAACAGCAGGAGCCCAACCGTACACGAGCAGATCACCAAG	5733						
Db	481	ACCTCCGCGAGAGCCCAACAGCAGGAGCCCAACCGTACACGAGCAGATCACCAAG	540						
QY	5734	CGCGTGGCTCCAGCCAGCGCGCCGCGAAGCCGCCAGCACCCTCCGAGAGCCAGCACA	5793						
Db	541	CGCGTGGCTCCAGCCAGCGCGCCAGCAGGAGCCGCCAGCACCCTCCGAGAGCCAGCACA	600						
QY	5794	CCCCACCGCTACC-----GCAGGGGCGGACCGAGCTGCGGAGGGAAGTCTCTCTGGC	5847						
Db	601	CGCCACCGCTACCAGGAGAGAGAGGCTCGGACAGAGCTGCGGAGGGAAGTCTCTGGC	660						
QY	5848	CGCCCTCTGGAGAGAGAGTCCCGCGCGATGCTCAGCAGCGGAGAGAGCGGTCC	5907						
Db	661	CGCCCTCTGGAGAGAGAGTCCCGCGCGATGCTCAGCAGCGGAGAGAGCGGTCC	719						
QY	5908	CCCGG 5912							
Db	720	CAGGG 724							

RESULT 14  
BM904785  
LOCUS  
DEFINITION  
AGENCY 6699306 NIH\_MGC\_72 Homo sapiens cdna clone IMAGE:5557590  
5', mRNA Sequence.  
ACCESSION  
BM904785  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1085)  
NIH-MGC <http://mgi.nhl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Tissue Procurement: ATCC/DCTD/DTF  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12279 row: d column: 07  
High quality sequence stop: 510.  
Location/Qualifiers  
1. .1085  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5557590"  
/issue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/notes="Organ: Skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 9.2%; Score 566.4; DB 12; Length 1085; Best Local Similarity 97.4%; Pred. No. 5.2e-123; Matches 629; Conservative 0; Mismatches 11; Indels 6; Gaps 5;									
QY	867	GGCGGTGATTCCTATGAGATGATTTATGGAGATCCCCCTCCAGAGGAAACCTCTGC	926						
Db	1	GGCGGTGATTCCTATGAGATGATTTATGGAGATCCCCCTCCAGAGGAAACCTCTGC	60						
QY	927	CAGAACCTTCAATTAACATTAATGAAATTCAGCGGTTTTGAAATTTCCAGATGACCCAA	986						
Db	61	CAGAACCTTCAATTAACATTAATGAAATTCAGCGGTTTTGAAATTTCCAGATGACCCAA	120						
QY	987	AGTGAGCAGTGACTTCTTGATCTGATTCAGAGTCTGTCGGCCAGAAAGAGAGACT	1046						
Db	121	AGTGAGCAGTGACTTCTTGATCTGATTCAGAGTCTGTCGGCCAGAAAGAGAGACT	180						
QY	1047	GAAGTTTGAAGGCTTTTGCTGCCATCCTTTCTCTTAAATTTGACTGGAAACAACTTCG	1106						
Db	181	GAAGTTTGAAGGCTTTTGCTGCCATCCTTTCTCTTAAATTTGACTGGAAACAACTTCG	240						
QY	1107	TAACCTCTCTCCCTCTGTTCCACCTCAAGCTGAGAGTACACCTCCCAATTTGA	1166						
Db	241	TAACCTCTCTCCCTCTGTTCCACCTCAAGCTGAGAGTACACCTCCCAATTTGA	300						
QY	1167	TGAACAGAGAGAAATTCGTTGGGTTTCATCTCTCGTGCAGCTGAGCCCTCAGGCTT	1226						
Db	301	TGAACAGAGAGAAATTCGTTGGGTTTCATCTCTCGTGCAGCTGAGCCCTCAGGCTT	360						
QY	1227	CTCGGTGAAGAACTGCGGTTTGTGGGTTTTCGTACAGAAAGCACTGGGATTCCTGG	1286						
Db	361	CTCGGTGAAGAACTGCGGTTTGTGGGTTTTCGTACAGAAAGCACTGGGATTCCTGG	420						
QY	1287	TAGATCTGAGTCTGTTGTTGCTGGGTCTGGACTCCCTGCCAAGACTAGCTCCATGAAA	1346						
Db	421	TAGATCTGAGTCTGTTGTTGCTGGGTCTGGACTCCCTGCCAAGACTAGCTCCATGAAA	480						
QY	1347	GAAGACTTCTCATCAAAA-GCAAAGAGTACAAGACTCTCAGGCAAGTGTACAAAGATG	1405						
Db	481	GAAGACTTCTCATCAAAAAGCAAGAGTACCAGACTCTCAGGCAAGTGTACAAAGATG	540						
QY	1406	AGCAGGAATGAGCC--GGTTACATCGAGAGTGTACAGGTGGAGCTGTG-CTTAGTC	1462						
Db	541	AGCAGGAATGAGCCCGGGTTACATCCGAAGTGTACAGGTGGAGCTGTGCTTAGTC	600						
QY	1463	AGAAGGAGGTGG-AGCTGAAGGCT-CTGAGACTCAGAGATCCCTC 1506							
Db	601	ACAAGAGGGGGAAGCTGAAGGCTCTCTGAGACTCAGAACATCCCC 646							

RESULT 15  
 B0807302  
 LOCUS  
 DEFINITION B0807302 640 bp mRNA linear EST 31-JUL-2002  
 NISC\_k001b10.y1 NCI\_CGAP\_Brn72 Macaca mulatta cDNA clone  
 IMAGE:5330250 5', mRNA sequence.  
 ACCESSION B0807302  
 VERSION B0807302.1 GI:22031511  
 KEYWORDS EST.  
 SOURCE Macaca mulatta (rhesus monkey)  
 ORGANISM Macaca mulatta  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.  
 REFERENCE 1 (bases 1 to 640)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
 cDNA Library Preparation:  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Plate: LLAM11838 row: C column: 19  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1..640  
 /organism="Macaca mulatta"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9544"  
 /clone="IMAGE:5330250"  
 /tissue\_type="hypothalamus"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Brn72"  
 /note="Organ: Brain; Vector: pCMV-Sport6.1; Site 1: NotI;  
 Site 2: EcoRV; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2.2 Kb. Constructed by Invitrogen.  
 Note: this is a NCI\_CGAP Library."

## FEATURES

source

## ORIGIN

Query Match 9.0%; Score 556.8; DB 13; Length 640;  
 Best Local Similarity 98.8%; Pred. No. 6.9e-121;  
 Matches 561; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4070 GCCTGTGCCCCCGCATCCAGCCGACAGAGAGTCTCAACTCCAGAGGAATTAGTC 4129  
 Db 1 GCCTGTGCCCCCGCATCCAGCCGACAGAGAGTCTCAACTCCAGAGGAATTAGTC 60  
 QY 4130 GCGCTTTAAGGAACGCGATGACCAACAATATTCCTCACCGATTCAACGTAGGACTGAACA 4189  
 Db 61 GCGCTTTAAGGAACGCGATGACCAACAATATTCCTCACCGATTCAACGTAGGACTGAACA 120  
 QY 4190 TCGGAGCCACAAGTGTGCTGTCTGGATACCGTGCACCTTTGGACGCCAGGCATCCA 4249  
 Db 121 TCGGAGCCACAAGTGTGCTGTCTGGATACCGTGCACCTTTGGACGCCAGGCATCCA 180  
 QY 4250 AATGTCTCGAATGTCTAGTGTGTCTACCCCAAGTGTCCACGTGCTTGCAGCCACCT 4309  
 Db 181 AATGTCTCGAATGTCTAGTGTGTCTACCCCAAGTGTCCACGTGCTTGCAGCCACCT 240  
 QY 4310 GCGGCTTGCGTCTGAATATGCGCACACACTTCAACGAGGCTTCTGCGTGACAAAATGA 4369  
 Db 241 GCGGCTTGCGTCTGCGAATACGCGCACACACTTCACTGAGGCTTCTGCGTGACAAAATGA 300  
 QY 4370 ACTCCCCAGGCTCTCAGACCAAGAGGCCAGCAGCAGCTTGCACCTGGAAGGGTGGATGA 4429  
 Db 301 ACTCCCCAGGCTCTCAGACCAAGAGGCCAGCAGCAGCTTGCACCTGGAAGGGTGGATGA 360  
 QY 4430 AGGTGCCCGAGGAATAACAAACAGGACAGCAAGGCTGGGACAGGAAGTACATTCTCTCTGG 4489

Db 361 AGGTGCCCGAGGAATAACAAACAGGACAGCAAGCTGGGACAGGAAGTACATTCTCTGG 420  
 QY 4490 AGGGATCAAAAAGTCTCTCATTTATGCAATGAAGCCAGAGAAGCTGGACAGAGGCCGGTGG 4549  
 Db 421 AGGGATCAAAAAGTCTCTCATTTATGCAATGAAGCCAGAGAAGCTGGACAGAGGCCGGTGG 480  
 QY 4550 AAGAATTTGAGCTGTGCTTCCCGACCGGGATGTATCTATTCATGTGCGGTTGGTGCCTT 4609  
 Db 481 AAGAATTTGAGCTGTGCTTCCCGACCGGGATGTATCTATTCATGTGCGGTTGGTGCCTT 540  
 QY 4610 CCGAACTCGCAAAATACAGCCAAAGCAGA 4637  
 Db 541 CCGAACTCGCAAAATACAGCCAAAGCAGA 568

Search completed: July 3, 2004, 19:29:05  
 Job time : 9474 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 16:52:20 ; Search time 1715 Seconds  
(without alignments)  
17324.953 Million cell updates/sec

Title: US-10-017-216-3  
Perfect score: 6159  
Sequence: 1 atgttgaagtccaatatgg.....ttctgagaacagattatgc 6159

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3163042 seqs, 2412103800 residues

Total number of hits satisfying chosen parameters: 6326084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6159	100.0	6159	14	US-10-017-216-3
2	6159	100.0	6162	15	US-10-325-430-11
3	6159	100.0	6574	14	US-10-017-216-1
4	6159	100.0	6574	15	US-10-325-430-10
5	5666.4	92.0	6298	13	US-10-415-011-43
6	5661.4	91.9	6165	14	US-10-028-946-1
7	5631	91.4	6189	13	US-10-262-511-1
8	5629.8	91.4	6201	13	US-09-964-956-10
9	5629.8	91.4	6201	13	US-10-262-511-13
10	5629.8	91.4	6201	13	US-09-964-956-8
11	5358	87.0	5877	14	US-10-028-946-3
12	2564.8	41.6	3131	13	US-10-276-774-137
13	2414	39.2	2542	13	US-10-262-511-7
14	2317.2	37.6	2497	13	US-10-262-511-5

15	1841	29.9	1870	13	US-10-262-511-3	Sequence 3, Appli
16	1524.2	24.7	1870	13	US-10-262-511-9	Sequence 9, Appli
17	1467.6	23.8	1915	13	US-10-262-511-11	Sequence 11, Appl
18	1393.4	22.6	2066	17	US-10-311-034-33	Sequence 33, Appl
19	1391.8	22.6	1515	9	US-09-804-471A-1	Sequence 1, Appli
20	1391.8	22.6	1515	15	US-10-238-709-1	Sequence 1, Appli
21	1391.8	22.6	1515	17	US-10-724-594-1	Sequence 1, Appli
22	952.4	15.5	2162	16	US-10-120-988-419	Sequence 419, App
23	950.6	15.4	957	16	US-10-413-897-1	Sequence 1, Appli
24	750	12.2	1133	15	US-10-282-048-1	Sequence 1, Appli
25	701.4	11.4	995	13	US-10-276-774-255	Sequence 255, App
26	421.4	6.8	2870	13	US-10-423-114-28241	Sequence 26241, A
27	350.4	5.7	354	14	US-10-040-739-9	Sequence 9, Appli
28	258	4.2	258	16	US-10-305-720-513	Sequence 513, App
29	243.6	4.0	1530	13	US-10-415-011-24	Sequence 24, Appl
30	243.6	4.0	4698	17	US-10-703-496-1	Sequence 1, Appli
31	243.6	4.0	5373	13	US-10-333-314-40	Sequence 40, Appl
32	236.8	3.8	6335	16	US-10-388-934-85	Sequence 85, Appl
33	234	3.8	5373	17	US-10-702-496-5	Sequence 5, Appli
34	228.2	3.7	2785	13	US-10-342-887-930	Sequence 930, App
35	228.2	3.7	2785	13	US-10-172-118-930	Sequence 26, Appl
36	228.2	3.7	5694	16	US-10-363-892-26	Sequence 26, Appl
37	228.2	3.7	7151	17	US-10-288-798-26	Sequence 39, Appl
38	228.2	3.7	2621	15	US-10-433-794-39	Sequence 179, App
39	225	3.7	5347	16	US-10-252-157-179	Sequence 99, Appl
40	221.6	3.6	5347	16	US-10-210-130-99	Sequence 101, App
41	221.6	3.6	5875	16	US-10-210-130-101	Sequence 1, Appli
42	220.8	3.6	3407	9	US-09-971-845-1	Sequence 9, Appli
43	220.8	3.6	3407	12	US-10-380-235-9	Sequence 1449, Ap
44	208.4	3.4	3363	16	US-10-012-697-1449	Sequence 9, Appli
45	208.4	3.4	6780	12	US-10-399-225-9	

ALIGNMENTS

RESULT 1  
US-10-017-216-3  
; Sequence 3, Application US/10017216  
; Publication No. US20020160483A1  
; GENERAL INFORMATION:  
; APPLICANT: KAPELIER-LIBERMANN, Rosana  
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prote  
; TITLE OF INVENTION: Kinase and Uses Therefor  
; FILE REFERENCE: 10147-57U1  
; CURRENT APPLICATION NUMBER: US/10/017,216  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: US 60/242,429  
; PRIOR FILING DATE: 2000-10-23  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 6159  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-216-3

Query Match 100.0%; Score 6159; DB 14; Length 6159;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGTTGAAGTTCAATAATGAGCGCGGAATCCTTTTGGATCGTGGTGTGCTGAACCCATT 60  
Db 1 ATGTTGAAGTTCAATAATGAGCGCGGAATCCTTTTGGATCGTGGTGTGCTGAACCCATT 60  
QY 61 GCCAGCGGGCTCCAGGCTGAATCTCTTTCAGGGGAAACACCCCTTATGACTCAA 120  
Db 61 GCCAGCGGGCTCCAGGCTGAATCTCTTTCAGGGGAAACACCCCTTATGACTCAA 120  
QY 121 CAGCAGATGCTCTCTCTTTCCGAGAGGATATTAGATCCCTCTTTGTTCTCTTTGAA 180  
Db 121 CAGCAGATGCTCTCTCTTTCCGAGAGGATATTAGATCCCTCTTTGTTCTCTTTGAA 180

181 QY GAATGCAGTCAGCCTGCTCTGATGAAGATTAGCACGTCGAGCAACTTTGTCCGGAGATAT 240  
181 Db GAATGCAGTCAGCCTGCTCTGATGAAGATTAGCACGTCGAGCAACTTTGTCCGGAGATAT 240  
241 QY TCCGACACCATAGCTAGCTGAGTTTACAGGAGCTCCAGCCTTCGCGAAAGGACATTCGAAGTCAGA 300  
241 Db TCCGACACCATAGCTAGCTGAGTTTACAGGAGCTCCAGCCTTCGCGAAAGGACATTCGAAGTCAGA 300  
301 QY AGCTCTGTAGTTGTGGTCACTTTGTCTGAAGTCAGGTGTTAAGAGAGAAAGCAACCGGG 360  
301 Db AGCTCTGTAGTTGTGGTCACTTTGTCTGAAGTCAGGTGTTAAGAGAGAAAGCAACCGGG 360  
361 QY GACATCTATCTATGAAGTGAAGAGAGAGAGGCTTTATGGCCACAGGACAGGTTTCA 420  
361 Db GACATCTATCTATGAAGTGAAGAGAGAGGCTTTATGGCCACAGGACAGGTTTCA 420  
421 QY TTTTGTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480  
421 Db TTTTGTGAGGAAGAGCGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480  
481 QY CAGTATGCCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540  
481 Db CAGTATGCCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540  
541 QY GACTTGTCTGCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAATACTGATACAGTTT 600  
541 Db GACTTGTCTGCTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAATACTGATACAGTTT 600  
601 QY TACTAGCTAGCTGATTTTGGCTGTTTCACAGGTTTCATCTGATGGGATACGTGATCGA 660  
601 Db TACTAGCTAGCTGATTTTGGCTGTTTCACAGGTTTCATCTGATGGGATACGTGATCGA 660  
661 QY GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGCTGATTTT 720  
661 Db GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGCTGATTTT 720  
721 QY GATCTGCCGCGAAAATGAATTCAAAACAGATGGTGAATGCCAAACTCCCGATTGGACC 780  
721 Db GATCTGCCGCGAAAATGAATTCAAAACAGATGGTGAATGCCAAACTCCCGATTGGACC 780  
781 QY CCAGATTACATGCTCCTCAAGTGTGACTGTGATCAACGGGATGGAAGGACCTTAC 840  
781 Db CCAGATTACATGCTCCTCAAGTGTGACTGTGATCAACGGGATGGAAGGACCTTAC 840  
841 QY GGCCTGGACTGTGACTGTGGTGGTCAAGTGGCGTGAATGCGCTATGAGATGATTTATGGAGA 900  
841 Db GGCCTGGACTGTGACTGTGGTGGTCAAGTGGCGTGAATGCGCTATGAGATGATTTATGGAGA 900  
901 QY TCCGCCCTCCGAGGAGACCTTCCAGAACCTTCAATAACATTTATGAAATTCAGCGG 960  
901 Db TCCGCCCTCCGAGGAGACCTTCCAGAACCTTCAATAACATTTATGAAATTCAGCGG 960  
961 QY TTTTGTGAAATTTCCAGATGACCCCAAGTGACAGTGAATTTCTTTGATCTGATTCAAAAGC 1020  
961 Db TTTTGTGAAATTTCCAGATGACCCCAAGTGACAGTGAATTTCTTTGATCTGATTCAAAAGC 1020  
1021 QY TTGTTGTGGCGCAGAAAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTCTTTCTTC 1080  
1021 Db TTGTTGTGGCGCAGAAAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTCTTTCTTC 1080  
1081 QY TCTAAAATTTGACTGGAACAACATTTGTTAACTCTCTCTCCCTCCCTTCTGTTCCACCTCAAG 1140  
1081 Db TCTAAAATTTGACTGGAACAACATTTGTTAACTCTCTCTCCCTCCCTTCTGTTCCACCTCAAG 1140  
1141 QY TCTGAGATGACCTCCAAATTTTATGAACACAGAGAGAAATTCGTGGGTTTCATCTCT 1200  
1141 Db TCTGAGATGACCTCCAAATTTTATGAACACAGAGAGAAATTCGTGGGTTTCATCTCT 1200  
1201 QY CGTGCAGCTCAGCCCTCAGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTTTCG 1260  
1201 Db CGTGCAGCTCAGCCCTCAGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTTTCG 1260  
1261 QY TACAGCAAGGCACCTGGGATTTCTGGTAGATCTGTAGTCTGTGTGTGGGCTCGGACTCC 1320

1261 Db TACAGCAAGGCACCTGGGATTTCTGGTAGATCTGTGTGTGGGCTCGGACTCC 1320  
1321 QY CCTGCCAAGACTAGCTCCATCGAAAGAAACTTTCTCATCAAAAGCAAGAGACTACAAGAC 1380  
1321 Db CCTGCCAAGACTAGCTCCATCGAAAGAAACTTTCTCATCAAAAGCAAGAGACTACAAGAC 1380  
1381 QY TCTCAGACAAGTGTCTCAAGATGGAGCAGAAATGACCCGGTTTACATCGAGAGTGTCA 1440  
1381 Db TCTCAGACAAGTGTCTCAAGATGGAGCAGAAATGACCCGGTTTACATCGAGAGTGTCA 1440  
1441 QY GAGTGTGAGGCTGTGCTTAGTCAGAAAGAGGTGAGGCTGAAGGCTCTGAGACTCAGAGA 1500  
1441 Db GAGTGTGAGGCTGTGCTTAGTCAGAAAGAGGTGAGGCTGAAGGCTCTGAGACTCAGAGA 1500  
1501 QY TCCCTCTCTGGAGCAGGACCTTGTCTTACCTACATCAAGATGTCAGTAGCTTTAAAGCGAAGT 1560  
1501 Db TCCCTCTCTGGAGCAGGACCTTGTCTTACCTACATCAAGATGTCAGTAGCTTTAAAGCGAAGT 1560  
1561 QY TTGGAGCAAGCAGCGATGGAGGTGTCCAGAGAGATGACAAAGCAGCTGAGCTTCTCCAT 1620  
1561 Db TTGGAGCAAGCAGCGATGGAGGTGTCCAGAGAGATGACAAAGCAGCTGAGCTTCTCCAT 1620  
1621 QY GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAGCAGGAGTAGTACCAGGCTCAA 1680  
1621 Db GATATCAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAGCAGGAGTAGTACCAGGCTCAA 1680  
1681 QY GTGGAAGAAATGAGTTGATGATGATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740  
1681 Db GTGGAAGAAATGAGTTGATGATGATCAGTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740  
1741 QY CGGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800  
1741 Db CGGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATCAAG 1800  
1801 QY CGGAAGCGCAGAAATGTCAGCAATAAATGTTGAAGGCTAAGGATCAAGGGAAGCCGTGAA 1860  
1801 Db CGGAAGCGCAGAAATGTCAGCAATAAATGTTGAAGGCTAAGGATCAAGGGAAGCCGTGAA 1860  
1861 QY GTGGAGAAATTCGAAACTGGAGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
1861 Db GTGGAGAAATTCGAAACTGGAGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
1921 QY CTCAAGAGAAACTGGAGAGGCTGCAAAAGGAGCGAGAGGAGGCTGGAGAAAGCTG 1980  
1921 Db CTCAAGAGAAACTGGAGAGGCTGCAAAAGGAGCGAGAGGAGGCTGGAGAAAGCTG 1980  
1981 QY CAGAACCGAGAGATTTCTTCTGAAGCATCAGAAAGAGCTGGAAGCTGGAAGCTGAGGAACGC 2040  
1981 Db CAGAACCGAGAGATTTCTTCTGAAGCATCAGAAAGAGCTGGAAGCTGGAAGCTGAGGAACGC 2040  
2041 QY CGCCATTTCTCTGGAGAACCAAGGTAAGAGACTAGAGACCATGAGCGCTAGAGAAAACAGA 2100  
2041 Db CGCCATTTCTCTGGAGAACCAAGGTAAGAGACTAGAGACCATGAGCGCTAGAGAAAACAGA 2100  
2101 QY CTGAAGATGACATCCAGACAAATCCCAAGTCTCAGCCCGACACCTAGAGTGCACCTG 2160  
2101 Db CTGAAGATGACATCCAGACAAATCCCAAGTCTCAGCCCGACACCTAGAGTGCACCTG 2160  
2161 QY GAGCTCGAAGAGAAAACATCGGAGGCGCCCAAGTCTCAGCCCGACACCTAGAGTGCACCTG 2220  
2161 Db GAGCTCGAAGAGAAAACATCGGAGGCGCCCAAGTCTCAGCCCGACACCTAGAGTGCACCTG 2220  
2221 QY AAACAGAAAGAGCAGCACTTATGAGGAAAAGATTAAAGTGTGGACAAATCAGATAAAGAAA 2280  
2221 Db AAACAGAAAGAGCAGCACTTATGAGGAAAAGATTAAAGTGTGGACAAATCAGATAAAGAAA 2280  
2281 QY GACCTGCTGACAAAGAGACACTGGAGAACATGATCGACAGACACGAGGAGAGGCCCCAT 2340  
2281 Db GACCTGCTGACAAAGAGACACTGGAGAACATGATCGACAGACACGAGGAGAGGCCCCAT 2340  
2341 QY GAGAAAGGCAAAATTCCTCAGCGAACAGAAAGGCGATGATCAATGCTATGATTTCCAGATC 2400

Db 2341 GAGAAAGGCAAAATCTCAGCGAACAGAGGGGATGATCAATGCTATGGATTCCAAGATC 2400  
QY AGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAATAAATCTTCAGCAAAATAGC 2460  
Db AGATCCCTGGAAACAGAGGATTTGGAACTGTCTGAAGCCAATAAATCTTCAGCAAAATAGC 2460  
QY AGTCCTTTTACCCAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520  
Db AGTCCTTTTACCCAAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAA 2520  
QY CAGAAATTTTACTCTGGAGACACAGGCTGGAGTTGGAGGCCAGAACCGGAAACTGGAG 2580  
Db CAGAAATTTTACTCTGGAGACACAGGCTGGAGTTGGAGGCCAGAACCGGAAACTGGAG 2580  
QY GAGCAGCTGGAGAAATCAGCCACCAAGACACACAGTGAAGAATCGCTGCTGGAATG 2640  
Db GAGCAGCTGGAGAAATCAGCCACCAAGACACACAGTGAAGAATCGCTGCTGGAATG 2640  
QY GAGCAAGATTCGGGAGGTGAGTCTAGAGCAGGAGGACAGAACTGGAGCTCAAGGCG 2700  
Db GAGCAAGATTCGGGAGGTGAGTCTAGAGCAGGAGGACAGAACTGGAGCTCAAGGCG 2700  
QY CAGCTCACAGAGCTACAGCTCTCCCTGAGAGGCGGAGTCAAGTTGACAGCCCTGGAG 2760  
Db CAGCTCACAGAGCTACAGCTCTCCCTGAGAGGCGGAGTCAAGTTGACAGCCCTGGAG 2760  
QY GCTGACGCGGCGCCCTGGAGAGCCAGCTTCCGAGGCGAAGACAGAGCTGGAAGAGCC 2820  
Db GCTGACGCGGCGCCCTGGAGAGCCAGCTTCCGAGGCGAAGACAGAGCTGGAAGAGCC 2820  
QY ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAATCCAGGCG 2880  
Db ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCACGGCACATAGAGATGAATCCAGGCG 2880  
QY AAATTTGATGCTCTTCGTAACAGCTGTACTGTAATCACAGACCTGGAGGACAGCTAAAC 2940  
Db AAATTTGATGCTCTTCGTAACAGCTGTACTGTAATCACAGACCTGGAGGAGCAGCTAAAC 2940  
QY CAGCTGACCGAGGACAAAGCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTC 3000  
Db CAGCTGACCGAGGACAAAGCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTC 3000  
QY GATGAGGCTTCGGGCGCAACGACAGAGATTGACAACTGCGAAGTGAAGTGACCATCTC 3060  
Db GATGAGGCTTCGGGCGCAACGACAGAGATTGACAACTGCGAAGTGAAGTGACCATCTC 3060  
QY CGCGGGAGATCACGGAACGAGAGATGACGTTACAGCCAGAAACAAAGATGAGGCT 3120  
Db CGCGGGAGATCACGGAACGAGAGATGACGTTACAGCCAGAAACAAAGATGAGGCT 3120  
QY CTGAAGACCACTGACCACTGCTGGAGAACAGGTCTATGGATTTGGAGGCCCTAAACGAT 3180  
Db CTGAAGACCACTGACCACTGCTGGAGAACAGGTCTATGGATTTGGAGGCCCTAAACGAT 3180  
QY GAGCTGCTAGAAAAGAGCGGCACTGGAGGCGCTGGAGAGCGTCTGGGTGATGAGAAA 3240  
Db GAGCTGCTAGAAAAGAGCGGCACTGGAGGCGCTGGAGAGCGTCTGGGTGATGAGAAA 3240  
QY TCCAGATTTGAGTGTGCGGTTCCAGAGCTGCAGAGATGCTGGACACCGAGAAAACAGAGC 3300  
Db TCCAGATTTGAGTGTGCGGTTCCAGAGCTGCAGAGATGCTGGACACCGAGAAAACAGAGC 3300  
QY AGGGCAGAGCCGATCAGCGGATTCACCGAGTCTCCGAGGTTGGAGTGGAGTGAAG 3360  
Db AGGGCAGAGCCGATCAGCGGATTCACCGAGTCTCCGAGGTTGGAGTGGAGTGAAG 3360  
QY GAGCAACAGGCTGAGATTTCTGCTCTGACAGAGGCTCTCAAGAGCAGAACTGAAGGCC 3420  
Db GAGCAACAGGCTGAGATTTCTGCTCTGACAGAGGCTCTCAAGAGCAGAACTGAAGGCC 3420  
QY GAGAGCTCTCTCACAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAAT 3480  
Db GAGAGCTCTCTCACAAGCTCAATGACCTGGAGAGAGAGCATGCTATGCTTGAATGAAT 3480

QY GCCCGAAGCTTACAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3540  
Db GCCCGAAGCTTACAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAA 3540  
QY GAGCAAGCCAAATTTACAGCAGCAGATGGACCTCGAGAAAATCACAATTTTCCGTCTGACT 3600  
Db GAGCAAGCCAAATTTACAGCAGCAGATGGACCTCGAGAAAATCACAATTTTCCGTCTGACT 3600  
QY CAAGGACTGCAAGAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAAGTGAAGTTCG 3660  
Db CAAGGACTGCAAGAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAAGTGAAGTTCG 3660  
QY GAGTATCAGCTGCAAAAACATTCAGAGTTCTCTATTTCTCATGAAAAGTGAATAATGGAAGC 3720  
Db GAGTATCAGCTGCAAAAACATTCAGAGTTCTCTATTTCTCATGAAAAGTGAATAATGGAAGC 3720  
QY ACTATTTCTCAACAAACCAAACTCATTTCTGCAAGCCAAATGGAACCACTGCT 3780  
Db ACTATTTCTCAACAAACCAAACTCATTTCTGCAAGCCAAATGGAACCACTGCT 3780  
QY AAAAGAAAAGGTTTATTTAGTCGACGAAAGAGGACCTCTTTACCCACACAGGTT 3840  
Db AAAAGAAAAGGTTTATTTAGTCGACGAAAGAGGACCTCTTTACCCACACAGGTT 3840  
QY CCTCTGAGTACAAATGAGCTGAAAGCTGGCCCTCGAAGAGGAGAAAGCTCTGTGCAGAG 3900  
Db CCTCTGAGTACAAATGAGCTGAAAGCTGGCCCTCGAAGAGGAGAAAGCTCTGTGCAGAG 3900  
QY CTAGAGAAAGCCCTTCAAGAGACCGGATCGAGCTCGGTCGCGCCGCGGAGGAGCTGCC 3960  
Db CTAGAGAAAGCCCTTCAAGAGACCGGATCGAGCTCGGTCGCGCCGCGGAGGAGCTGCC 3960  
QY CACGCAAAACCAAGGACCAACCCACATCCACGCGACCCACCGGACCGGAGGAGCAGATC 4020  
Db CACGCAAAACCAAGGACCAACCCACATCCACGCGACCCACCGGAGGAGCAGATC 4020  
QY GCCATGTCGCCATTCGTCGCGTCCAGAGACACAGCCCAAGTGAACATGCGAGGCCACA 4080  
Db GCCATGTCGCCATTCGTCGCGTCCAGAGACACAGCCCAAGTGAACATGCGAGGCCACA 4080  
QY CCGCATCCAGCGCAGAAAAGGAGTCTCAACTCCAGAGGAAATTTAGTCGGCGCTTTAAG 4140  
Db CCGCATCCAGCGCAGAAAAGGAGTCTCAACTCCAGAGGAAATTTAGTCGGCGCTTTAAG 4140  
QY GAACGATGACCAACAATATTCCTCAACCGATTCAAGTAGGACTGAACATGCGAGGCCACA 4200  
Db GAACGATGACCAACAATATTCCTCAACCGATTCAAGTAGGACTGAACATGCGAGGCCACA 4200  
QY AAGTGTCTGTGTCTGGATACCGTGCACTTTGGAGCCAGGCAATCCAATGCTCTGAA 4260  
Db AAGTGTCTGTGTCTGGATACCGTGCACTTTGGAGCCAGGCAATCCAATGCTCTGAA 4260  
QY TGTGAGTGTGTGTCAACCCAAAGTGTCCACGTCGTTGCGAGCCACCTCGGCGTTGCCCT 4320  
Db TGTGAGTGTGTGTCAACCCAAAGTGTCCACGTCGTTGCGAGCCACCTCGGCGTTGCCCT 4320  
QY GCTGAATATGCCACACACTTCAACCGAGCTTCTGCGTGAACAAATGAATCTCCCGAGGT 4380  
Db GCTGAATATGCCACACACTTCAACCGAGCTTCTGCGTGAACAAATGAATCTCCCGAGGT 4380  
QY CTCAGACCAAGGAGCCAGCAGCAGCTTCACCTGGAAGGGTGGATGAAGGTGCCCAGG 4440  
Db CTCAGACCAAGGAGCCAGCAGCAGCTTCACCTGGAAGGGTGGATGAAGGTGCCCAGG 4440  
QY AATAACAAACGAGGACAGCAAGCTGGGACAGGAAGTACATTTCTCTGGAGGATCAAAA 4500  
Db AATAACAAACGAGGACAGCAAGCTGGGACAGGAAGTACATTTCTCTGGAGGATCAAAA 4500  
QY GTCCTCATTTATGAATGAAGCCAGAGAGCTGGACAGAGCGCGGTGGAGGATTTGAG 4560  
Db GTCCTCATTTATGAATGAAGCCAGAGAGCTGGACAGAGCGCGGTGGAGGATTTGAG 4560

4561 CTGTGCTTCCCGACGGGATGTATCTATTATCATGTCGCGTGGTGGTCTTCCGAATCTCGCA 4620  
4561 CTGTGCTTCCCGACGGGATGTATCTATTATCATGTCGCGTGGTGGTCTTCCGAATCTCGCA 4620  
4621 AATACAGCCAAAGCAGAAAAGCAGAAAGCTGATGCTTAAACTGTCTTGGAAACTCCCTGCTG 4680  
4621 AATACAGCCAAAGCAGAAAAGCAGAAAGCTGATGCTTAAACTGTCTTGGAAACTCCCTGCTG 4680  
4681 AACTGGAAGTGATGACCGTCTAGACATGAACCTGACCGCTGCCCTTCAAGTACACAGGTG 4740  
4681 AACTGGAAGTGATGACCGTCTAGACATGAACCTGACCGCTGCCCTTCAAGTACACAGGTG 4740  
4741 GTGTGTTGGGGACCCAGAGGAGGCTCTACGCGCTGAATGTCTTGAAGAACTCCCTTAACC 4800  
4741 GTGTGTTGGGGACCCAGAGGAGGCTCTACGCGCTGAATGTCTTGAAGAACTCCCTTAACC 4800  
4801 CATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATTAATCAAGGACCTGGAGAAGCTA 4860  
4801 CATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATTAATCAAGGACCTGGAGAAGCTA 4860  
4861 CTGATGATAGCAGGAAGAGCGGCACTGTGCTTGTGGAGCTGAAGAAAGTGAACAG 4920  
4861 CTGATGATAGCAGGAAGAGCGGCACTGTGCTTGTGGAGCTGAAGAAAGTGAACAG 4920  
4921 TCCTGGCCGAGTCCCACTTCCTGCGCCAGCCGACATCTCACCCAACTTTTGAAGCT 4980  
4921 TCCTGGCCGAGTCCCACTTCCTGCGCCAGCCGACATCTCACCCAACTTTTGAAGCT 4980  
4981 GTCAAGGGTGCACATTTGTTGGGGCAGGCAAGATTGAGAAAGGGCTCTGCACTGTGCA 5040  
4981 GTCAAGGGTGCACATTTGTTGGGGCAGGCAAGATTGAGAAAGGGCTCTGCACTGTGCA 5040  
5041 GCCATGCCAGCAAGTCTGCTTCTCCGCTACACGAAACCTCAGCAAACTACTGCATC 5100  
5041 GCCATGCCAGCAAGTCTGCTTCTCCGCTACACGAAACCTCAGCAAACTACTGCATC 5100  
5101 CGAAAGATAGAGACCTCAGAGCCCTGAGCTGATATCCATCTTCAACAAATACAGTATC 5160  
5101 CGAAAGATAGAGACCTCAGAGCCCTGAGCTGATATCCATCTTCAACAAATACAGTATC 5160  
5161 CTCATTGGAAACCAATAAATTTCTACGAAATCGCATGAAGCAGTACACGCTCGAGGAATTC 5220  
5161 CTCATTGGAAACCAATAAATTTCTACGAAATCGCATGAAGCAGTACACGCTCGAGGAATTC 5220  
5221 CTGATTAAGAAATGACCAATTCCTTGGCACTGCTGTGTTTTCGCGCTCTTCCAAACAGCTTC 5280  
5221 CTGATTAAGAAATGACCAATTCCTTGGCACTGCTGTGTTTTCGCGCTCTTCCAAACAGCTTC 5280  
5281 CCTGTCTCAATCGTGAGGTACACGCGCAGGCGAGGAGGAGTACTTGTGCTGTGTTTC 5340  
5281 CCTGTCTCAATCGTGAGGTACACGCGCAGGCGAGGAGGAGTACTTGTGCTGTGTTTC 5340  
5341 CAGAAATTTGGAGTGTCTGGAATCTTACGGAAGACGTAGCGGCACAGACGATCTCAAG 5400  
5341 CAGAAATTTGGAGTGTCTGGAATCTTACGGAAGACGTAGCGGCACAGACGATCTCAAG 5400  
5401 TGGAGTCGCTTACCTTTGGCTTTGCTTACAGAGAACCTTATCTGTTGTGACCCACTTC 5460  
5401 TGGAGTCGCTTACCTTTGGCTTTGCTTACAGAGAACCTTATCTGTTGTGACCCACTTC 5460  
5461 AACTCACTCGAAGTAATGAGATCCAGGACCGCTCCTCAGCAGGAGCCCTCGCCGAGCG 5520  
5461 AACTCACTCGAAGTAATGAGATCCAGGACCGCTCCTCAGCAGGAGCCCTCGCCGAGCG 5520  
5521 TACTGGACATCCGAAACCCGCGCTACTTGGGCGCTGCGCATTTCTCAGGAGCGATTTAC 5580  
5521 TACTGGACATCCGAAACCCGCGCTACTTGGGCGCTGCGCATTTCTCAGGAGCGATTTAC 5580  
5581 TTGGCGTCTCTATACAGAGTAATTTAAGGGTCATTTGCTGCAAGGAAACCTCGTGAAG 5640  
5581 TTGGCGTCTCTATACAGAGTAATTTAAGGGTCATTTGCTGCAAGGAAACCTCGTGAAG 5640  
5641 GAGTCGGCACTGAACACACCCGGGCGCGTCCACTCCCGCAGCAGCCCCCAACAGCGA 5700

Db 5641 GAGTCGGCACTGAACACACACCGGGCGCGTCCACTCCCGCAGCAGCCCCCAACAGCGA 5700  
Qy 5701 GGCCCAACCCAGTAAACAGAGACATCAACAGCGCGTGGCTCCAGCCAGCGCCGCC 5760  
Db 5701 GGCCCAACCCAGTAAACAGAGACATCAACAGCGCGTGGCTCCAGCCAGCGCCGCC 5760  
Qy 5761 GAAGCCCAACCCAGTAAACAGAGCGCAAGCAGACACCCCGCTACCGAGGGCGGAC 5820  
Db 5761 GAAGCCCAACCCAGTAAACAGAGCGCAAGCAGACACCCCGCTACCGAGGGCGGAC 5820  
Qy 5821 GAGTCGCGAGGGAACAAGTCTCTGCGCCCGCTGAGCGAGAGAGTCCCCCGCGCG 5880  
Db 5821 GAGTCGCGAGGGAACAAGTCTCTGCGCCCGCTGAGCGAGAGAGTCCCCCGCGCG 5880  
Qy 5881 ATGCTCAGCAGCGGAGAGAGCGTCCCGCGGAGGCTGTTGAAAGACAGCAGGGGC 5940  
Db 5881 ATGCTCAGCAGCGGAGAGAGCGTCCCGCGGAGGCTGTTGAAAGACAGCAGGGGC 5940  
Qy 5941 CGGCTGCTCGGAGCGCTGAGACCCCGCTGTCCAGGTGAACAGGGAAGAGGCGAG 6000  
Db 5941 CGGCTGCTCGGAGCGCTGAGACCCCGCTGTCCAGGTGAACAGGGAAGAGGCGAG 6000  
Qy 6001 AGTCCCTCTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTT 6060  
Db 6001 AGTCCCTCTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTT 6060  
Qy 6061 GACAACTGCGCAGTAACTGCTGAGTAACTGCTGAGTAACTGCTGAGTAACTGCTGAGTAACT 6120  
Db 6061 GACAACTGCGCAGTAACTGCTGAGTAACTGCTGAGTAACTGCTGAGTAACTGCTGAGTAACT 6120  
Qy 6121 CAGCAGTTGAAAAGTCTGTTCTGAGAACAGATTATTC 6159  
Db 6121 CAGCAGTTGAAAAGTCTGTTCTGAGAACAGATTATTC 6159

RESULT 2  
US-10-325-430-11  
; Sequence 11, Application US/10325430  
; Publication No. US2003015325A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc  
; APPLICANT: Silos-Santiago, Inmaculada  
; APPLICANT: Rosenfeld, Julie Beth  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING  
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,  
; FILE REFERENCE: MPI01-294P.RNM  
; CURRENT APPLICATION NUMBER: US/10/325,430  
; PRIOR FILING DATE: 2002-12-19  
; PRIOR APPLICATION NUMBER: US 60/341,953  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 6162  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(6162)  
US-10-325-430-11

Query Match 100.0%; Score 6159; DB 15; Length 6162;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGTAAGTTCAAATATGAGCGCGGAAATCTTTGGATGCTGCTGCTGTAACCAAT 60  
Db 1 ATGTGTAAGTTCAAATATGAGCGCGGAAATCTTTGGATGCTGCTGCTGTAACCAAT 60  
Qy 61 GCACCGCGGCTCCAGGCTGAATCTGTTCTCAGGGAAACCACTTTATGACTCAA 120

Db 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120  
Qy 121 CAGCAGATGTCCTCTCTTTCCGAGAAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
Db 121 CAGCAGATGTCCTCTCTTTCCGAGAAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
Qy 181 GAATCAGTCAGCTGCTCTGATGAAAGATTAAAGCAAGTCAGCAACTTTGTCGGAAGTAT 240  
Db 181 GAATCAGTCAGCTGCTCTGATGAAGATTAAAGCAAGTCAGCAACTTTGTCGGAAGTAT 240  
Qy 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGGCTTCGGCAAGGACTTCGAAAGTCAGA 300  
Db 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGGCTTCGGCAAGGACTTCGAAAGTCAGA 300  
Qy 301 AGTCTTTAGTGTGCTGCTCTGCTGAAAGTCAGGTCGAGTGGTAAGAGAAAGCAACCGGG 360  
Db 301 AGTCTTTAGTGTGCTGCTCTGCTGAAAGTCAGGTCGAGTGGTAAGAGAAAGCAACCGGG 360  
Qy 361 GACATCTATGCTATGAAAGTGAAGAGAAAGGCTTTATTGGCCCGAGGACAGTTTCA 420  
Db 361 GACATCTATGCTATGAAAGTGAAGAGAAAGGCTTTATTGGCCCGAGGACAGTTTCA 420  
Qy 421 TTTTGTGAGGAGAGCGGAAACATATTATCTGAAACCAAGCCCGTGGATCCCCAAATTA 480  
Db 421 TTTTGTGAGGAGAGCGGAAACATATTATCTGAAACCAAGCCCGTGGATCCCCAAATTA 480  
Qy 481 CAGTATGCTTTTCAGGACAAATAATCACCTTTATCTGATGGAGGATATCAGCCTGGAGG 540  
Db 481 CAGTATGCTTTTCAGGACAAATAATCACCTTTATCTGATGGAGGATATCAGCCTGGAGG 540  
Qy 541 GACTTGTCTGCTCTTTGAAATAGATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600  
Db 541 GACTTGTCTGCTCTTTGAAATAGATGAGGACCAAGTTAGATGAAACCTGATACAGTTT 600  
Qy 601 TACCTAGCTGAGCTGATTTGGCTGTTTACAGGTTCTATCTGATGGGATACGTCATCGA 660  
Db 601 TACCTAGCTGAGCTGATTTGGCTGTTTACAGGTTCTATCTGATGGGATACGTCATCGA 660  
Qy 661 GACATCAGCCTGAGAACATTTCTGTTGACCCACAGGACATCAAGCTGGTGAATTT 720  
Db 661 GACATCAGCCTGAGAACATTTCTGTTGACCCACAGGACATCAAGCTGGTGAATTT 720  
Qy 721 GGATCTGCCCGGAAATGAATTCAAACAGATGTTGAATGCCAACTCCCGATTGGGACC 780  
Db 721 GGATCTGCCCGGAAATGAATTCAAACAGATGTTGAATGCCAACTCCCGATTGGGACC 780  
Qy 781 CCAGATTACATGGCTCCTGAAGTGTGCTGATGAACCGGGATGGAAGGACCTTAC 840  
Db 781 CCAGATTACATGGCTCCTGAAGTGTGCTGATGAACCGGGATGGAAGGACCTTAC 840  
Qy 841 GGCCTGGACTGTGACTGGTGGTCAAGTGGGCGTGAATTCATGAGATGATTTATGGGAGA 900  
Db 841 GGCCTGGACTGTGACTGGTGGTCAAGTGGGCGTGAATTCATGAGATGATTTATGGGAGA 900  
Qy 901 TCCCCCTTCGAGAGGGAACCTTCGCGAGAACCTTCAATTAACATTAATTAATTCAGCGG 960  
Db 901 TCCCCCTTCGCGAGGGAACCTTCGCGAGAACCTTCAATTAACATTAATTAATTCAGCGG 960  
Qy 961 TTTTTCGAAATTTCCAGATGACCCCAAGTGAGCAGTGAATTTCTGATCTGATTCGAAAGC 1020  
Db 961 TTTTTCGAAATTTCCAGATGACCCCAAGTGAGCAGTGAATTTCTGATCTGATTCGAAAGC 1020  
Qy 1021 TTGTTGTGGGCGCAGAAAGAGAGACTGAAAGTTTGAAGTTCCTTGGTCCCATCTTTCTTC 1080  
Db 1021 TTGTTGTGGGCGCAGAAAGAGAGACTGAAAGTTTGAAGTTCCTTGGTCCCATCTTTCTTC 1080  
Qy 1081 TCTAAAATTGACTGGAAACACATTCGTAATCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1140  
Db 1081 TCTAAAATTGACTGGAAACACATTCGTAATCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1140  
Qy 1141 TCTGACGATGACACCTCCCAATTTTGAATGAACACAGAGAAATTCGTGGGTTTCATCTCT 1200  
Db 1141 TCTGACGATGACACCTCCCAATTTTGAATGAACACAGAGAAATTCGTGGGTTTCATCTCT 1200

Qy 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGFTGAAGAACTGCGCTTTGTGGGTTTTTCG 1260  
Db 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGFTGAAGAACTGCGCTTTGTGGGTTTTTCG 1260  
Qy 1261 TACAGCAAGCACTGGGATCTTTGTAGATCTGAGTCTGTTGTTGTCGGGCTTGGACTCC 1320  
Db 1261 TACAGCAAGCACTGGGATCTTTGTAGATCTGAGTCTGTTGTTGTCGGGCTTGGACTCC 1320  
Qy 1321 CTTGCCAAGACTAGCTCCATGGAAAAAACTTCTCATCAAAAAGCAAGAGCTTACAGAC 1380  
Db 1321 CTTGCCAAGACTAGCTCCATGGAAAAAACTTCTCATCAAAAAGCAAGAGCTTACAGAC 1380  
Qy 1381 TCTCAGCAAGTGTCAAGATGGAGCAAGAAATGACCCGTTTACATCGAAGTGTCA 1440  
Db 1381 TCTCAGCAAGTGTCAAGATGGAGCAAGAAATGACCCGTTTACATCGAAGTGTCA 1440  
Qy 1441 GAGGTGAGGCTGTGCTTAGTCAGAAAGGAGTGGAGCTGAAGGCTTCTGAGACTCAGAGA 1500  
Db 1441 GAGGTGAGGCTGTGCTTAGTCAGAAAGGAGTGGAGCTGAAGGCTTCTGAGACTCAGAGA 1500  
Qy 1501 TCCCTCTGAGCAGGACCTTGTCTACTACATCAGAAATGCAAGGCTTCTCCAT 1560  
Db 1501 TCCCTCTGAGCAGGACCTTGTCTACTACATCAGAAATGCAAGGCTTCTCCAT 1560  
Qy 1561 TTGGAGCAAGCAAGGAGTGTCCAGAGGATGCAAAAAGCACTGCAAGCTTCTCCAT 1620  
Db 1561 TTGGAGCAAGCAAGGAGTGTCCAGAGGATGCAAAAAGCACTGCAAGCTTCTCCAT 1620  
Qy 1621 GATATCAGAGCAGAGCCGAAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGGCTCAA 1680  
Db 1621 GATATCAGAGCAGAGCCGAAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGGCTCAA 1680  
Qy 1681 GTGGAGAAATGAGTTGATGATCACTTGGAGAGGATCTTGTCTCAGCAAGAAAGA 1740  
Db 1681 GTGGAGAAATGAGTTGATGATCACTTGGAGAGGATCTTGTCTCAGCAAGAAAGA 1740  
Qy 1741 CCGAGTCTCTACGAATCTGAGCTGAGAGACTCTCGGCTTCTGCTGAAGAAATTCAG 1800  
Db 1741 CCGAGTCTCTACGAATCTGAGCTGAGAGACTCTCGGCTTCTGCTGAAGAAATTCAG 1800  
Qy 1801 CGGAAACCGACAGAAATGTCAGCAATACTTTGAAGCTAAGGATCAAGGAGGACCTGAA 1860  
Db 1801 CGGAAACCGACAGAAATGTCAGCAATACTTTGAAGCTAAGGATCAAGGAGGACCTGAA 1860  
Qy 1861 GTGGGAAATATGCAAACTGGAGAAAGTCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920  
Db 1861 GTGGGAAATATGCAAACTGGAGAAAGTCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920  
Qy 1921 CTCGAGAGAACTGGAGAAAGCTGCAAGAGGAGCCGAGAGGAGCTGGAGAACTG 1980  
Db 1921 CTCGAGAGAACTGGAGAAAGCTGCAAGAGGAGCCGAGAGGAGCTGGAGAACTG 1980  
Qy 1981 CAGAACCGAGAGATTTCTTTGAAGGATCAGAAAGAGTGGTGGAGCTGAGAACTG 2040  
Db 1981 CAGAACCGAGAGATTTCTTTGAAGGATCAGAAAGAGTGGTGGAGCTGAGAACTG 2040  
Qy 2041 CGCCATTTCTTGGAGAACTAGAGACTAGAGACCTAGAGCCTAGAGAAACAGA 2100  
Db 2041 CGCCATTTCTTGGAGAACTAGAGACTAGAGACCTAGAGCCTAGAGAAACAGA 2100  
Qy 2101 CTGAAGATGACATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTG 2160  
Db 2101 CTGAAGATGACATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATCTG 2160  
Qy 2161 GAGCTCGAAGAAACATCGGAGGCCCCAAGTCTCAGCCCCAGCAGCTTGAAGTGCACCTG 2220  
Db 2161 GAGCTCGAAGAAACATCGGAGGCCCCAAGTCTCAGCCCCAGCAGCTTGAAGTGCACCTG 2220  
Qy 2221 AAAAGAGAGCAGCAGCTATGAGGAAAGAAATTAAGTGTGGCAATCAGATAAAGAAA 2280  
Db 2221 AAAAGAGAGCAGCAGCTATGAGGAAAGAAATTAAGTGTGGCAATCAGATAAAGAAA 2280

Qy	2281	GACCTGGCTGACAAGGAGACACTCGAGAAACATGATGCAGAGACACGAGGAGGAGGCCCAT	2340
Db	2281	GACCTGGCTGACAAGGAGACACTCGAGAAACATGATGCAGAGACACGAGGAGGAGGCCCAT	2340
Qy	2341	GAGAAGGGCAAAATTTCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCACAAGATC	2400
Db	2341	GAGAAGGGCAAAATTTCTCAGCGAACAGAGGCGATGATCAATGCTATGGATTCACAAGATC	2400
Qy	2401	AGATCCCTTGGAAACAGAGATTTGGAACTGTTGAAGCCAAATAAACTTGCAGCAAAATAGC	2460
Db	2401	AGATCCCTTGGAAACAGAGATTTGGAACTGTTGAAGCCAAATAAACTTGCAGCAAAATAGC	2460
Qy	2461	AGTCTTTTTTACCACCAAGGAAACATCAAGGCCCAACAGAGAGATGATTTCTGAACTCAGGCAA	2520
Db	2461	AGTCTTTTTTACCACCAAGGAAACATCAAGGCCCAACAGAGAGATGATTTCTGAACTCAGGCAA	2520
Qy	2521	CAGAAATTTTACCCTTGGAGACACAGCGCTGGGAAGTTGGAGGCCCAAGAACCCGAAACTGGAG	2580
Db	2521	CAGAAATTTTACCCTTGGAGACACAGCGCTGGGAAGTTGGAGGCCCAAGAACCCGAAACTGGAG	2580
Qy	2581	GAGCAGCTGGAGAGATCAGCCACCAAGACCAAGTGCAGAGAAATCGGCTGCTGGAACTG	2640
Db	2581	GAGCAGCTGGAGAGATCAGCCACCAAGACCAAGTGCAGAGAAATCGGCTGCTGGAACTG	2640
Qy	2641	GAGACAAGATTTCGGGGAGGTCAGTCTAGAGCACGAGGAGCAGAAATCTGGAGCTCAAGCGC	2700
Db	2641	GAGACAAGATTTCGGGGAGGTCAGTCTAGAGCACGAGGAGCAGAAATCTGGAGCTCAAGCGC	2700
Qy	2701	CAGCTCACAGAGTACAGCTCTCCCTCGAGAGCGGAGTCAAGTTGACAGCCCTGCAG	2760
Db	2701	CAGCTCACAGAGTACAGCTCTCCCTCGAGAGCGGAGTCAAGTTGACAGCCCTGCAG	2760
Qy	2761	GCTCCACGGCGGCCCTTGGAGAGCCAGCTTCGCCAGCGAAGACAGAGCTTGGAAAGAGACC	2820
Db	2761	GCTCCACGGCGGCCCTTGGAGAGCCAGCTTCGCCAGCGAAGACAGAGCTTGGAAAGAGACC	2820
Qy	2821	ACAGCAGAACTGAAGAGGAGATCCAGGCACCTCACGGCACATAGAGATGAAATCCAGCGC	2880
Db	2821	ACAGCAGAACTGAAGAGGAGATCCAGGCACCTCACGGCACATAGAGATGAAATCCAGCGC	2880
Qy	2881	AAATTTGATGCTCTTCGTACAGCTGTACTGTAAATCACAGACCTGGAGGAGAGCTAAAC	2940
Db	2881	AAATTTGATGCTCTTCGTACAGCTGTACTGTAAATCACAGACCTGGAGGAGAGCTAAAC	2940
Qy	2941	CAGCTGACCGAGGACAAACGCTGAACCTCAACAAACCAAACTTCTACTTGTCCAAACAACTC	3000
Db	2941	CAGCTGACCGAGGACAAACGCTGAACCTCAACAAACCAAACTTCTACTTGTCCAAACAACTC	3000
Qy	3001	GATGAGGCTTTGGCGCCAAACGAGAGATTGTACAACTGCGAAGTCAAGTGGACCAATCTC	3060
Db	3001	GATGAGGCTTTGGCGCCAAACGAGAGATTGTACAACTGCGAAGTCAAGTGGACCAATCTC	3060
Qy	3061	CGCCGGGAGATCACGGAAACGAGAGATGCAGCTTACCAGCCAGAAACAAACGATGGAGGCT	3120
Db	3061	CGCCGGGAGATCACGGAAACGAGAGATGCAGCTTACCAGCCAGAAACAAACGATGGAGGCT	3120
Qy	3121	CTGAAGACCAAGTGCACCAATCTTGGAGGAACAGGTCAATGGAATTTGGAGGCCCTTAAACGAT	3180
Db	3121	CTGAAGACCAAGTGCACCAATCTTGGAGGAACAGGTCAATGGAATTTGGAGGCCCTTAAACGAT	3180
Qy	3181	GAGCTGCTAGAAAAAGACGGCAGTGGAGGCTTGGAGGAGCGCTCTGGGTGATGAGAAA	3240
Db	3181	GAGCTGCTAGAAAAAGACGGCAGTGGAGGCTTGGAGGAGCGCTCTGGGTGATGAGAAA	3240
Qy	3241	TCCACGTTTCAGTGTCCGGTTTCAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGC	3300
Db	3241	TCCACGTTTCAGTGTCCGGTTTCAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGC	3300
Qy	3301	AGGGCAGAGCCGATCAGCGGATCACCGAGTCTCCAGGTGGTGAGCTGCCAGTGAAG	3360
Db	3301	AGGGCAGAGCCGATCAGCGGATCACCGAGTCTCCAGGTGGTGAGCTGCCAGTGAAG	3360
Yy	3361	GAGCACAAGGCTCAGATTTCTCGCTCTGACAGCGGCTCTCAAAAGACAGAAGCTGAAGGCC	3420

Db	3361	GAGCACAAGGCTGAGATTCTCGCTCTCGACGAGCTCTCAAAGCAGCAGAAGCTTGAAGGCC	3420
Qy	3421	GAGAGCCTCTCTGACAAGCTCAATGACCTGGAGAAGAAGCATGCTATGCTTGAATGAAT	3480
Db	3421	GAGAGCCTCTCTGACAAGCTCAATGACCTGGAGAAGAAGCATGCTATGCTTGAATGAAT	3480
Qy	3481	GCCCGAAGCTTTACAGCAGAAGCTGGAGACTGAAACGAGAGCTCAAACAGAGGCTTCTGGAA	3540
Db	3481	GCCCGAAGCTTTACAGCAGAAGCTGGAGACTGAAACGAGAGCTCAAACAGAGGCTTCTGGAA	3540
Qy	3541	GAGCAAGCCTTTACAGCAGCAGATGGACCTGCAGAAAAATCACATTTTCCGCTCTGACT	3600
Db	3541	GAGCAAGCCTTTACAGCAGCAGATGGACCTGCAGAAAAATCACATTTTCCGCTCTGACT	3600
Qy	3601	CAAGGACTGCAAGAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAAAGTGAAGTCTTG	3660
Db	3601	CAAGGACTGCAAGAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAAAGTGAAGTCTTG	3660
Qy	3661	GAGTATCAGCTGGAAAAATTCAGGTTCTCTATTCTCATGAAGAGTGAATAATGGAAGGC	3720
Db	3661	GAGTATCAGCTGGAAAAATTCAGGTTCTCTATTCTCATGAAGAGTGAATAATGGAAGGC	3720
Qy	3721	ACTATTCTCAACAAAAACCAACTCAATGATTTTCTGCAAGCCAAAAATGGAACCAACTGCT	3780
Db	3721	ACTATTCTCAACAAAAACCAACTCAATGATTTTCTGCAAGCCAAAAATGGAACCAACTGCT	3780
Qy	3781	AAAAAGAAAAAGGTTTATTATTGTCAGCGAAGAGAGAACCTGCTTTTACCCACACAGGTT	3840
Db	3781	AAAAAGAAAAAGGTTTATTATTGTCAGCGAAGAGAGAACCTGCTTTTACCCACACAGGTT	3840
Qy	3841	CCTCTGCACTACATGAGCTTGAAGCTGGCCCTGGAGAGGAGAGAAAGCTCGCTGTGCAGAG	3900
Db	3841	CCTCTGCACTACATGAGCTTGAAGCTGGCCCTGGAGAGGAGAGAAAGCTCGCTGTGCAGAG	3900
Qy	3901	CTAGAGAAAGCCTTTCAGAGAACCGGATCGAGCTCCGGTCCGCCCGGAGAGAGCTGCC	3960
Db	3901	CTAGAGAAAGCCTTTCAGAGAACCGGATCGAGCTCCGGTCCGCCCGGAGAGAGCTGCC	3960
Qy	3961	CACCGCAAGCAACGGACACCCACACCCATCCACGCCAGCCACCGCAGGAGCAGAGATC	4020
Db	3961	CACCGCAAGCAACGGACACCCACACCCATCCACGCCAGCCACCGCAGGAGCAGAGATC	4020
Qy	4021	GCCATGTCGCCATCGTCCGCTCCAGAGACCCAGCCAGTGCCATGAGCCTGTCTGGCC	4080
Db	4021	GCCATGTCGCCATCGTCCGCTCCAGAGACCCAGCCAGTGCCATGAGCCTGTCTGGCC	4080
Qy	4081	CCGCCATCAGCCGACAGAGAGTCTTCAATCCAGAGGAATTTAGTCCGGCTCTTAAG	4140
Db	4081	CCGCCATCAGCCGACAGAGAGTCTTCAATCCAGAGGAATTTAGTCCGGCTCTTAAG	4140
Qy	4141	GAACCGATGCACACAATATTCCTCACCGATTCAACGTAGGACTGAACATGGAGCCACA	4200
Db	4141	GAACCGATGCACACAATATTCCTCACCGATTCAACGTAGGACTGAACATGGAGCCACA	4200
Qy	4201	AAGTGTCTGTGTCTGGATACCGTGCATTTGGACGCCAGGCATCCAAATGTCTCGAA	4260
Db	4201	AAGTGTCTGTGTCTGGATACCGTGCATTTGGACGCCAGGCATCCAAATGTCTCGAA	4260
Qy	4261	TGTCAGGTGATGTGTCAACCCCAAGTGTCTCCAGTGTCTGCCAGCCACTTCGGCTTGCT	4320
Db	4261	TGTCAGGTGATGTGTCAACCCCAAGTGTCTCCAGTGTCTGCCAGCCACTTCGGCTTGCT	4320
Qy	4321	GCTGAATATGCCACACATTTACCGAGGCTTCTGCCGTGACAAATGAACCTCCACAGT	4380
Db	4321	GCTGAATATGCCACACATTTACCGAGGCTTCTGCCGTGACAAATGAACCTCCACAGT	4380
Qy	4381	CTCCAGACCAAGAGCCACAGCAGCTTGCACCTCGAAGGTTGGATGAAGTGCACAG	4440
Db	4381	CTCCAGACCAAGAGCCACAGCAGCTTGCACCTCGAAGGTTGGATGAAGTGCACAG	4440
Qy	4441	AATAACAAACGAGGACAGCAAGCTGGGACAGGAAGTACATTTGCTCGAGGGATCAAAA	4500

```
Db 4441 AATAACAAACGAGCAGCAGGCTGGGACAGGAAGTACATTGTCTCGAGGGATCAAAA 4500
QY 4501 GTCCTCATTTATGACAAATGAAGCCAGAGAAGCTGACAGAGGCGGTGGAGAAATTTGAG 4560
Db 4501 GTCCTCATTTATGACAAATGAAGCCAGAGAAGCTGACAGAGGCGGTGGAGAAATTTGAG 4560
QY 4561 CTGTGCTCTCCGACGGGGATGATCTATTTCATGTCGCGTGGTGGTTCGGAATCGCA 4620
Db 4561 CTGTGCTCTCCGACGGGGATGATCTATTTCATGTCGCGTGGTGGTTCGGAATCGCA 4620
QY 4621 AATACAGCCAAAGCAGAAAACGAGAGCTGATGCTAACTGCTTGGAAAATCCCTGCTG 4680
Db 4621 AATACAGCCAAAGCAGAAAACGAGAGCTGATGCTAACTGCTTGGAAAATCCCTGCTG 4680
QY 4681 AAACTGGAAGGTGATGACCGCTCTAGACATGAACCTGACGCTGCCCTTCAGTGACAGGTG 4740
Db 4681 AAACTGGAAGGTGATGACCGCTCTAGACATGAACCTGACGCTGCCCTTCAGTGACAGGTG 4740
QY 4741 GTGTTGGTGGGACCGAGGAAGGCTACGCGCTGAATGCTTGAATAATCCCTAAC 4800
Db 4741 GTGTTGGTGGGACCGAGGAAGGCTACGCGCTGAATGCTTGAATAATCCCTAAC 4800
QY 4801 CATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATCAAGGACCTGGAGAAGCTA 4860
Db 4801 CATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATCAAGGACCTGGAGAAGCTA 4860
QY 4861 CTCATGATAGAGGAGAGAGCGGCACTGTGCTCTTGTGGAGCTGAAGAAAGTGAACAG 4920
Db 4861 CTCATGATAGAGGAGAGAGCGGCACTGTGCTCTTGTGGAGCTGAAGAAAGTGAACAG 4920
QY 4921 TCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTACCCAAATTTTGAAGCT 4980
Db 4921 TCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTACCCAAATTTTGAAGCT 4980
QY 4981 GTCAAGGGCTGCCACTTGTTTGGGGCAGGCAAGATTGAGAAGCGCTCTGCATCTGTGCA 5040
Db 4981 GTCAAGGGCTGCCACTTGTTTGGGGCAGGCAAGATTGAGAAGCGCTCTGCATCTGTGCA 5040
QY 5041 GCCATGCCAGAAAGTCTGATCTTCGCTACAGGAAGCTCAGCAATACTGCAATC 5100
Db 5041 GCCATGCCAGAAAGTCTGATCTTCGCTACAGGAAGCTCAGCAATACTGCAATC 5100
QY 5101 CGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTACCATAATACAGTATC 5160
Db 5101 CGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTACCATAATACAGTATC 5160
QY 5161 CTCATTGGAACCAATAAATCTACGAATTCGACATGAAGCAGTACACGCTCAGGAATTC 5220
Db 5161 CTCATTGGAACCAATAAATCTACGAATTCGACATGAAGCAGTACACGCTCAGGAATTC 5220
QY 5221 CTGGATAAGATGACCATTCCTTGGCACCTGCTGTTTGGCGCTCTTCCAAACAGCTTC 5280
Db 5221 CTGGATAAGATGACCATTCCTTGGCACCTGCTGTTTGGCGCTCTTCCAAACAGCTTC 5280
QY 5281 CTTGCTCTCAATCGTCAGGTGAACAGCGAGGGCAGCAGAGAGTACTGCTGTGTTTC 5340
Db 5281 CTTGCTCTCAATCGTCAGGTGAACAGCGAGGGCAGCAGAGAGTACTGCTGTGTTTC 5340
QY 5341 CACGAATTTGGAGTGTCTGGAATCTTACGGAAGACGTAGCGGACAGACGATCTCAAG 5400
Db 5341 CACGAATTTGGAGTGTCTGGAATCTTACGGAAGACGTAGCGGACAGACGATCTCAAG 5400
QY 5401 TGGAGTCTGCTTACCTTTGGCCTTTGCTTACAGAAACCCCTATCTGTTGTGACCCACTTC 5460
Db 5401 TGGAGTCTGCTTACCTTTGGCCTTTGCTTACAGAAACCCCTATCTGTTGTGACCCACTTC 5460
QY 5461 AACTCATCGAAGTAATTGAGATCCAGGACCGCTCTCAGAGGGACCCCTGCCGAGCG 5520
Db 5461 AACTCATCGAAGTAATTGAGATCCAGGACCGCTCTCAGAGGGACCCCTGCCGAGCG 5520
QY 5521 TACCTGACATCCGACCGCTACCTGGGCCCTGCCATTTCTCAGAGGAGATTAC 5580
Db 5521 TACCTGACATCCCGAACCCGCTACCTGGGCCCTTGCCATTTCTCAGAGGAGATTAC 5580
```

```
QY 5581 TTGGCGTCTCTATACCAGGATAAAATTAAGGGTCATTTTGTGCAAGGGAAACCTCGTGAAG 5640
Db 5581 TTGGCGTCTCTATACCAGGATAAAATTAAGGGTCATTTTGTGCAAGGGAAACCTCGTGAAG 5640
QY 5641 GAGTCCGGCACTGAACACACCGGGCGCTCCACCTCCGCGAGCAGCCCAACAAGCGA 5700
Db 5641 GAGTCCGGCACTGAACACACCGGGCGCTCCACCTCCGCGAGCAGCCCAACAAGCGA 5700
QY 5701 GGCCCAACCCACGTACAAACGAGCAGCATCACAAACGCGTGGCCTCCAGCCAGCGCCGCC 5760
Db 5701 GGCCCAACCCACGTACAAACGAGCAGCATCACAAACGCGTGGCCTCCAGCCAGCGCCGCC 5760
QY 5761 GAAGCCCCCAGCCACCCGCGAGAGCCAGACACACCCACCGCTACCGGAGGGCGGAC 5820
Db 5761 GAAGCCCCCAGCCACCCGCGAGAGCCAGACACACCCACCGCTACCGGAGGGCGGAC 5820
QY 5821 GAGCTGGCAGGAGCAAGTCTCTGGCGGCCCTTGAGCGAGAGAGTCCCCCGGCCGG 5880
Db 5821 GAGCTGGCAGGAGCAAGTCTCTGGCGGCCCTTGAGCGAGAGAGTCCCCCGGCCGG 5880
QY 5881 ATGCTAGCAGCGGAGAGAGCGGTCCCGGAGAGCTGTTTGAAGCAGCAGCAGGGGC 5940
Db 5881 ATGCTAGCAGCGGAGAGAGCGGTCCCGGAGAGCTGTTTGAAGCAGCAGCAGGGGC 5940
QY 5941 CGGCTGCTCGCGAGCGGTGAGAGCCCGCTGTCCAGGTGAACAGGAGAGGCGAG 6000
Db 5941 CGGCTGCTCGCGAGCGGTGAGAGCCCGCTGTCCAGGTGAACAGGAGAGGCGAG 6000
QY 6001 AGTGCCTCTCAAGTTTTACGGTTTACACCTGTCACTTATGACTGGAATAAAAAGCTG 6060
Db 6001 AGTGCCTCTCAAGTTTTACGGTTTACACCTGTCACTTATGACTGGAATAAAAAGCTG 6060
QY 6061 GACAACTGCGCAGTAACTGTCAGTCTGAGGATCATCCAGCTGAATGAGAAATCCGG 6120
Db 6061 GACAACTGCGCAGTAACTGTCAGTCTGAGGATCATCCAGCTGAATGAGAAATCCGG 6120
QY 6121 CAGCAGGTTGAAAAGTCTGTTCTCAGAACAGATTATTGC 6159
Db 6121 CAGCAGGTTGAAAAGTCTGTTCTCAGAACAGATTATTGC 6159

RESULT 3
US-10-017-216-1
; Sequence 1, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPLELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prote
; FILE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-216-1

Query Match 100.0%; Score 6159; DB 14; Length 6574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTCAAAATGAGCGCGGAATCCTTTGGATGCTGCTGCTGTAACCCATT 60
Db 19 ATGTTGAAGTCAAAATGAGCGCGGAATCCTTTGGATGCTGCTGCTGTAACCCATT 78
QY 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCCACCTTTATCACTCAA 120
```

Db 79 GCCAGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGAAACCAACCCCTTTATGACTCAA 138  
Qy 121 CACAGATGTCCTCTCTTTCCGAGAAGGATATTAGATGCCCTCTTGTCTCTTTGAA 180  
Db 139 CAGCAGATGTCCTCTCTTTCCGAGAAGGATATTAGATGCCCTCTTGTCTCTTTGAA 198  
Qy 181 GAATGCACTCAGCTCTCTCATGAAGATTAAAGCACTGAGCAACTTTGTCGGAAGTAT 240  
Db 199 GAATGCACTCAGCTCTCTCATGAAGATTAAAGCACTGAGCAACTTTGTCGGAAGTAT 258  
Qy 241 TCGACACCAATAGTGAAGTTACAGGAGCTCCAGCTTTCCGCAAGGACTTCGAAGTCAGA 300  
Db 259 TCGACACCAATAGTGAAGTTACAGGAGCTCCAGCTTTCCGCAAGGACTTCGAAGTCAGA 318  
Qy 301 AGTCTTGTAGTGTGCTCACTTTGCTGAACTGAGTGTGTAAGAGAGAAAGCAACCGGG 360  
Db 319 AGTCTTGTAGTGTGCTCACTTTGCTGAACTGAGTGTGTAAGAGAGAAAGCAACCGGG 378  
Qy 361 GACATCTATGCTATGAAGTGAAGAGAAAGGCTTTATTGCGCCAGAGCAGGTTTCA 420  
Db 379 GACATCTATGCTATGAAGTGAAGAGAAAGGCTTTATTGCGCCAGAGCAGGTTTCA 438  
Qy 421 TTTTGTGAGAGAGCGGACATATTATCTCGAGCACAGCCCGTGATCCGCCAATTA 480  
Db 439 TTTTGTGAGAGAGCGGACATATTATCTCGAGCACAGCCCGTGATCCGCCAATTA 498  
Qy 481 CAGTATGCTTTTCAGGACAAATAACCTTTATCTGATGAGGAATATCAGCTCGAGGG 540  
Db 499 CAGTATGCTTTTCAGGACAAATAACCTTTATCTGATGAGGAATATCAGCTCGAGGG 558  
Qy 541 GACTTGTGTCATTTTGAATAGATATGAGGACAGTTAGATGAAGAAACCTGATACAGTT 600  
Db 559 GACTTGTGTCATTTTGAATAGATATGAGGACAGTTAGATGAAGAAACCTGATACAGTT 618  
Qy 601 TACTAGCTGAGCTGATTTGGCTGTTTCACAGGTTTCATCTGATGGGATACGTGATCGA 660  
Db 619 TACTAGCTGAGCTGATTTGGCTGTTTCACAGGTTTCATCTGATGGGATACGTGATCGA 678  
Qy 661 GACATCAAGCTGAGAACATTTCTGTTGACCGGACAGGACACATCAAGCTGTGATTTT 720  
Db 679 GACATCAAGCTGAGAACATTTCTGTTGACCGGACAGGACACATCAAGCTGTGATTTT 738  
Qy 721 GGATCTGCGCGAAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTGGACC 780  
Db 739 GGATCTGCGCGAAATGAATTCAAACAAGATGGTGAATGCCAACTCCCGATTGGACC 798  
Qy 781 CCAGATTACATGCTCTGAACTGCTGATGAACCGGGATGGAAGGACCTTAC 840  
Db 799 CCAGATTACATGCTCTGAACTGCTGATGAACCGGGATGGAAGGACCTTAC 858  
Qy 841 GGCTCGGATGTCAGTGTGTTGAGTGGCGGTGATTCCTATGAGTGAATTTATGGAGA 900  
Db 859 GGCTCGGATGTCAGTGTGTTGAGTGGCGGTGATTCCTATGAGTGAATTTATGGAGA 918  
Qy 901 TCCCTCTCGCAGAGGAACTCTGCGAAGCTTCAATAACATTAATGAATTTCCAGGG 960  
Db 919 TCCCTCTCGCAGAGGAACTCTGCGAAGCTTCAATAACATTAATGAATTTCCAGGG 978  
Qy 961 TTTTGAATTTCCAGATGACCCCAAGTGACAGTGAATTTCTGATCTGATTCAGAGC 1020  
Db 979 TTTTGAATTTCCAGATGACCCCAAGTGACAGTGAATTTCTGATCTGATTCAGAGC 1038  
Qy 1021 TTCTTGTGCGGCAGAAAGAGACTGAAAGTTTGAAGTTCCTTGTGCTGCATCTCTTCTTC 1080  
Db 1039 TTCTTGTGCGGCAGAAAGAGACTGAAAGTTTGAAGTTCCTTGTGCTGCATCTCTTCTTC 1098  
Qy 1081 TCTAAATTTGATGGAACAATTTGTTACTCTCTCTCCCTCTTCTGTTCCCAACCTCAAG 1140  
Db 1099 TCTAAATTTGATGGAACAATTTGTTACTCTCTCTCCCTCTTCTGTTCCCAACCTCAAG 1158  
Qy 1141 TCTGAGGATGACACCTTCCAAATTTTGAATGAACAGAGAAAGATTCGTGGGTTTCACTCTCT 1200  
Db 1159 TCTGAGGATGACACCTTCCAAATTTTGAATGAACAGAGAAAGATTCGTGGGTTTCACTCTCT 1218

Qy 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCCGTTTGTGGGCTTTTCG 1260  
Db 1219 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTGAAGAACTGCCGTTTGTGGGCTTTTCG 1278  
Qy 1261 TACAGCAAGGCACTGGGATTTCTTGGTAGATCTCAGTCTGTGTGTGCGGTCTCGACTCC 1320  
Db 1279 TACAGCAAGGCACTGGGATTTCTTGGTAGATCTCAGTCTGTGTGTGCGGTCTCGACTCC 1338  
Qy 1321 CCGTCCAGGACTAGCTCAGTGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTCAGAGAC 1380  
Db 1339 CCGTCCAGGACTAGCTCAGTGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTCAGAGAC 1398  
Qy 1381 TCTCAGCAAAAGTGTCAAGATGGAGCAAAATGACCCGGTTTACATCGAGAGTGTCA 1440  
Db 1399 TCTCAGCAAAAGTGTCAAGATGGAGCAAAATGACCCGGTTTACATCGAGAGTGTCA 1458  
Qy 1441 GAGTGGAGGCTGTGCTTAGTTCAGAGAGGAGTGAAGCTGAAGGCTCTGAGACTCAGAGA 1500  
Db 1459 GAGTGGAGGCTGTGCTTAGTTCAGAGAGGAGTGAAGCTGAAGGCTCTGAGACTCAGAGA 1518  
Qy 1501 TCCCTCTCGAGCAGGACCTTGTCTACCTACATCACAAGATGCAAGTCTTAAAGCGAAGT 1560  
Db 1519 TCCCTCTCGAGCAGGACCTTGTCTACCTACATCACAAGATGCAAGTCTTAAAGCGAAGT 1578  
Qy 1561 TTGAGCAAGCAGGATGGAGGTGTCCAGGAGGATGACAAAGCACTGCAAGTCTTCCAT 1620  
Db 1579 TTGAGCAAGCAGGATGGAGGTGTCCAGGAGGATGACAAAGCACTGCAAGTCTTCCAT 1638  
Qy 1621 GATATCAGAGCAGAGCGGAGCTTCCAGAAATCAAGAGCAGGATGACAGGCTCAA 1680  
Db 1639 GATATCAGAGCAGAGCGGAGCTTCCAGAAATCAAGAGCAGGATGACAGGCTCAA 1698  
Qy 1681 GTGGAAGAAATGAGTTGATGATGAATCAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGA 1740  
Db 1699 GTGGAAGAAATGAGTTGATGATGAATCAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGA 1758  
Qy 1741 CGGAGTATCTTACGAACTGAGCTGAGAGAGTCTCGGCTTCTGCTGAGAGATTTCAAG 1800  
Db 1759 CGGAGTATCTTACGAACTGAGCTGAGAGAGTCTCGGCTTCTGCTGAGAGATTTCAAG 1818  
Qy 1801 CGGAAACGCAAGATGTCAAGCTTAACTGTTGAAGCTAAGGATCAAGGGAAGCTGAA 1860  
Db 1819 CGGAAACGCAAGATGTCAAGCTTAACTGTTGAAGCTAAGGATCAAGGGAAGCTGAA 1878  
Qy 1861 GTGGAGAAATATGCAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
Db 1879 GTGGAGAAATATGCAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1938  
Qy 1921 CTCCAGAGAACTGGAGAGGCTGCAAAAGAGCGAGCGGAGAGGAGCTGGAAGAGCTG 1980  
Db 1939 CTCCAGAGAACTGGAGAGGCTGCAAAAGAGCGAGCGGAGAGGAGCTGGAAGAGCTG 1998  
Qy 1981 CAGAACCGAGAGGATTTCTTGAAGGCTCAGAAAGAAAGCTGTTGAAAGCTGAGAAAGC 2040  
Db 1999 CAGAACCGAGAGGATTTCTTGAAGGCTCAGAAAGAAAGCTGTTGAAAGCTGAGAAAGC 2058  
Qy 2041 CGCCATCTCTGAGACAGGTTAAAGAGCTAGAGCCATGAGAGCTGAGAGAAACAGA 2100  
Db 2059 CGCCATCTCTGAGAGAAAGGTTAAAGAGCTAGAGCCATGAGAGCTGAGAGAAACAGA 2118  
Qy 2101 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATATAATTCG 2160  
Db 2119 CTGAAGGATGACATCCAGACAAATCCCAACAGATCCAGAGATGGCTGATATAATTCG 2178  
Qy 2161 GAGCTCGAAGAGAAACATCGGAGGCGCAAGTCTCAGCCAGCACCCTAGAGTGACCTG 2220  
Db 2179 GAGCTCGAAGAGAAACATCGGAGGCGCAAGTCTCAGCCAGCACCCTAGAGTGACCTG 2238  
Qy 2221 AAACAGAAAGAGCAGCACTATGAGGAAAGATTTAAAGTGTGGACAAATCAGATAAAGAAA 2280  
Db 2239 AAACAGAAAGAGCAGCACTATGAGGAAAGATTTAAAGTGTGGACAAATCAGATAAAGAAA 2298





QY 61 GCCACGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACACCTTTATGACTCAA 120  
DB |||||  
QY 79 GCCACCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACACCTTTATGACTCAA 138  
DB |||||  
QY 121 CAGCAGATGTCTCTCTTTCCGAGAGGAGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180  
DB |||||  
QY 139 CAGCAGATGTCTCTCTTTCCGAGAGGAGATATTAGATGCCCTCTTTGTTCTCTTTGAA 198  
DB |||||  
QY 181 GAATCAGTCAAGCTGCTCTGATGAGATTAAGCAGCTGACCACTTTGTCCGGAAGTAT 240  
DB |||||  
QY 199 GAATCAGTCAAGCTGCTCTGATGAGATTAAGCAGCTGACCACTTTGTCCGGAAGTAT 258  
DB |||||  
QY 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAAAGGACTTCGAAAGTCAGA 300  
DB |||||  
QY 259 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAAAGGACTTCGAAAGTCAGA 318  
DB |||||  
QY 301 AGCTTTGAGTTGTGTTCACTTTCTGAAAGTGAGGTGGTAAAGAGAAACACCGGG 360  
DB |||||  
QY 319 AGCTTTGAGTTGTGTTCACTTTCTGAAAGTGAGGTGGTAAAGAGAAACACCGGG 378  
DB |||||  
QY 361 GACATCTATGCTATGAAAGTGATGAAGAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 420  
DB |||||  
QY 379 GACATCTATGCTATGAAAGTGATGAAGAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 438  
DB |||||  
QY 421 TTTTGTGAGGAAGAGCGGAACAATATTATCTCGAAGCACAAGCCCTGGATCCCCAATTA 480  
DB |||||  
QY 439 TTTTGTGAGGAAGAGCGGAACAATATTATCTCGAAGCACAAGCCCTGGATCCCCAATTA 498  
DB |||||  
QY 481 CAGTATGCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540  
DB |||||  
QY 499 CAGTATGCTTTTCAGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 558  
DB |||||  
QY 541 GACTTGCTCTCACTTTTGAATAGATATGAGGACCCAGTTAGATGAAAACCTGATACAGTTT 600  
DB |||||  
QY 559 GACTTGCTCTCACTTTTGAATAGATATGAGGACCCAGTTAGATGAAAACCTGATACAGTTT 618  
DB |||||  
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCTTTCACTGATGGGATAGCTGATCGA 660  
DB |||||  
QY 619 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCTTTCACTGATGGGATAGCTGATCGA 678  
DB |||||  
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCCACAGACACATCAAGCTGGTGATTTT 720  
DB |||||  
QY 679 GACATCAAGCCTGAGAACATTTCTGTTGACCCACAGACACATCAAGCTGGTGATTTT 738  
DB |||||  
QY 721 GGATCTGCCGGAATAATGAATTCAAACAAGATGGTGAATGCCAACTCCGATTTGGGACC 780  
DB |||||  
QY 739 GGATCTGCCGGAATAATGAATTCAAACAAGATGGTGAATGCCAACTCCGATTTGGGACC 798  
DB |||||  
QY 781 CCAGATTATCGCTCTGAAGTGCTGACTGTGATGAACGGGATGGAAGGACACCTTAC 840  
DB |||||  
QY 799 CCAGATTATCGCTCTGAAGTGCTGACTGTGATGAACGGGATGGAAGGACACCTTAC 858  
DB |||||  
QY 841 GGCCTGGAATGCTGACTGGTGGTCAGTGGGCGTGATTGCCCTATCAGATGATTTATGGGAGA 900  
DB |||||  
QY 859 GGCCTGGAATGCTGACTGGTGGTCAGTGGGCGTGATTGCCCTATCAGATGATTTATGGGAGA 918  
DB |||||  
QY 901 TCCCTCTTCGAGAGGAACTCTGCGAGAACTTCAATAAATATGAATTTCCAGCGG 960  
DB |||||  
QY 919 TCCCTCTTCGAGAGGAACTCTGCGAGAACTTCAATAAATATGAATTTCCAGCGG 978  
DB |||||  
QY 961 TTTTTCGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCAAAGC 1020  
DB |||||  
QY 979 TTTTTCGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCBAAGC 1038  
DB |||||  
QY 1021 TTGTTGTGGGCGCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGTCGCATCTTTCTTC 1080  
DB |||||  
QY 1039 TTGTTGTGGGCGCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGTGTCGCATCTTTCTTC 1098  
DB |||||  
QY 1081 TCTAAATTTGACTGGACACATTCGTTACTCTCTCCCTCCCTCTCGTTCCACCTCAAG 1140  
DB |||||  
QY 1099 TCTAAATTTGACTGGACACATTCGTTACTCTCTCCCTCCCTCTCGTTCCACCTCAAG 1158  
DB |||||

QY 1141 TCTGACGATGACACCTCCAATTTTGTATGAACCAAGAGAAATTCGTGGGTTTCATCTCTCT 1200  
DB |||||  
QY 1159 TCTGACGATGACACCTCCAATTTTGTATGAACCAAGAGAAATTCGTGGGTTTCATCTCTCT 1218  
DB |||||  
QY 1201 CCGTCCAGCTGAGCCCTCAGCCTTCGGGTGAAGAACTCGCGTTTGTGGGGTTTTCG 1260  
DB |||||  
QY 1219 CCGTCCAGCTGAGCCCTCAGCCTTCGGGTGAAGAACTCGCGTTTGTGGGGTTTTCG 1278  
DB |||||  
QY 1261 TACAGCAAGCACTGGGGATTTCTTGTAGATCTGAGTCTGTGTGTGTCGGGTTCGGACTCC 1320  
DB |||||  
QY 1279 TACAGCAAGCACTGGGGATTTCTTGTAGATCTGAGTCTGTGTGTGTCGGGTTCGGACTCC 1338  
DB |||||  
QY 1321 CTTGCCAAGACTAGCTTCATCGGAAAGAAACTTCTCATCAAAAGCAAAAGCTTACAGAC 1380  
DB |||||  
QY 1339 CTTGCCAAGACTAGCTTCATCGGAAAGAAACTTCTCATCAAAAGCAAAAGCTTACAGAC 1398  
DB |||||  
QY 1381 TCTCAGACAAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTTACATCGGAGAGTGTCA 1440  
DB |||||  
QY 1399 TCTCAGACAAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTTACATCGGAGAGTGTCA 1458  
DB |||||  
QY 1441 GAGTGGAGCTGTCTTATTAGTCAAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1500  
DB |||||  
QY 1459 GAGTGGAGCTGTCTTATTAGTCAAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1518  
DB |||||  
QY 1501 TCCCTCTGAGCAGGACCTTGTCTACCTACATCACAGAATGCAGTAGCTTAAAGCCGAAGT 1560  
DB |||||  
QY 1519 TCCCTCTGAGCAGGACCTTGTCTACCTACATCACAGAATGCAGTAGCTTAAAGCCGAAGT 1578  
DB |||||  
QY 1561 TTGAGCAAGCAAGGATGAGGTGTCCAGGAGGATGACAAAACACTGCAAGCTTCTCCAT 1620  
DB |||||  
QY 1579 TTGAGCAAGCAAGGATGAGGTGTCCAGGAGGATGACAAAACACTGCAAGCTTCTCCAT 1638  
DB |||||  
QY 1621 GATATCAGAGCAGAGCCGAGGACTCCAGAATCAAGAGCAGGAGTACCAGGCTCAA 1680  
DB |||||  
QY 1639 GATATCAGAGCAGAGCCGAGGACTCCAGAATCAAGAGCAGGAGTACCAGGCTCAA 1698  
DB |||||  
QY 1681 GTGCAAGAAATGAGTTGATGATCAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740  
DB |||||  
QY 1699 GTGCAAGAAATGAGTTGATGATCAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1758  
DB |||||  
QY 1741 CGGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGTTCAG 1800  
DB |||||  
QY 1759 CGGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGTTCAG 1818  
DB |||||  
QY 1801 CGGAAAGCGACAGAAATGTCAAGTAAACTGTGTAAGGCTAAGGATCAAGGGAAGCCCTGAA 1860  
DB |||||  
QY 1819 CGGAAAGCGACAGAAATGTCAAGTAAACTGTGTAAGGCTAAGGATCAAGGGAAGCCCTGAA 1878  
DB |||||  
QY 1861 GTGGAGAAATATGCAAACTGGAGAAATCAATGCTGAGCAGCAGCTCAAAAATTCAGAG 1920  
DB |||||  
QY 1879 GTGGAGAAATATGCAAACTGGAGAAATCAATGCTGAGCAGCAGCTCAAAAATTCAGAG 1938  
DB |||||  
QY 1921 CTCCAAGAGAAACTGGAGAAAGCTTCTGTAAGGCTCAGAAAGAGCTGGTGAAGCTGAGGAAACGC 2040  
DB |||||  
QY 1939 CTCCAAGAGAAACTGGAGAAAGCTTCTGTAAGGCTCAGAAAGAGCTGGTGAAGCTGAGGAAACGC 2058  
DB |||||  
QY 1981 CAGAAACCGAGAGGATTTCTTCTGAAGGCTCAGAAAGAGCTGGTGAAGCTGAGGAAACGC 2100  
DB |||||  
QY 1999 CAGAAACCGAGAGGATTTCTTCTGAAGGCTCAGAAAGAGCTGGTGAAGCTGAGGAAACGC 2118  
DB |||||  
QY 2041 CGCATTTCTTGGAGAAACAAGTTAAAGAGACTAGAGACCATGAGCCGTGAGGAAACAGA 2160  
DB |||||  
QY 2059 CGCATTTCTTGGAGAAACAAGTTAAAGAGACTAGAGACCATGAGCCGTGAGGAAACAGA 2178  
DB |||||  
QY 2101 CTGAGGATGATCTCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGT 2180  
DB |||||  
QY 2119 CTGAGGATGATCTCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGT 2198  
DB |||||  
QY 2161 GAGCTCGAAGAGAAACATTCGGAGGCGCCAAAGTCTCAGCCAGCAGCCTAGAAAGTGCACCTG 2220  
DB |||||  
QY 2179 GAGCTCGAAGAGAAACATTCGGAGGCGCCAAAGTCTCAGCCAGCAGCCTAGAAAGTGCACCTG 2238  
DB |||||  
QY 2221 AAAACGAAAGACGACACTATGAGGAAAGAGATTAAGTGTGTGGACAATCAGATTAAGAAA 2280  
DB |||||

Db	2239	AAACAGAAAGACGACGACCTATGAGGAAAAGATTAAAGTTGTTGGACAATTCAGATAAAGAAA	2299
QY	2281	GACCTGGCTGACAAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGAGCGCCCAT	2340
Db	2299	GACCTGGCTGACAAGGAGACACTGGAGAACATGATGACAGACACGAGGAGGAGCGCCCAT	2358
QY	2341	GAGAAAGGCAAAATTCCTCAGCAAACAGAAAGCGATGATCAATGCTATGGATTCCAAAGATC	2400
Db	2359	GAGAAAGGCAAAATTCCTCAGCAAACAGAAAGCGATGATCAATGCTATGGATTCCAAAGATC	2418
QY	2401	AGATCCCTGGAAACAGAGGATTGTGGAACTGTCTGAGGCCAATAAAGCTTGAGGCAAAATGAGC	2460
Db	2419	AGATCCCTGGAAACAGAGGATTGTGGAACTGTCTGAGGCCAATAAAGCTTGAGGCAAAATGAGC	2478
QY	2461	AGTCTTTTTTACCAGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA	2520
Db	2479	AGTCTTTTTTACCAGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAA	2538
QY	2521	CAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCCACAGAACCCGAAAACCTGGAG	2580
Db	2539	CAGAAATTTTACCTGGAGACACAGGCTGGGAAGTTGGAGGCCACAGAACCCGAAAACCTGGAG	2598
QY	2581	GAGCAGCTGGGAAGATCAGGCCACCAAGACACACAGTCAAGAAATCGGCTGCTGGAACCTG	2640
Db	2599	GAGCAGCTGGGAAGATCAGGCCACCAAGACACACAGTCAAGAAATCGGCTGCTGGAACCTG	2658
QY	2641	GAGACAAATTTGGGGAGGTCAGTCTAGAGCACGAGGAGCAGAAATCGGAGCTCAAGCGC	2700
Db	2659	GAGACAAATTTGGGGAGGTCAGTCTAGAGCACGAGGAGCAGAAATCGGAGCTCAAGCGC	2718
QY	2701	CAGCTTCACAGAGCTACAGCTCTCCCTGCAGGAGCGCGAGTCACAGTTTGACAGCCCTGCAG	2760
Db	2719	CAGCTTCACAGAGCTACAGCTCTCCCTGCAGGAGCGCGAGTCACAGTTTGACAGCCCTGCAG	2778
QY	2761	GCTGCACGGGGGCCCTGGAGAGCCAGCTTCGCCAGCGGGAAGACAGAGCTGGGAAGAGACC	2820
Db	2779	GCTGCACGGGGGCCCTGGAGAGCCAGCTTCGCCAGCGGGAAGACAGAGCTGGGAAGAGACC	2838
QY	2821	ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGC	2880
Db	2839	ACAGCAGAAGCTGAAGAGGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGC	2898
QY	2881	AAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC	2940
Db	2899	AAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCAGCTAAAC	2958
QY	2941	CAGCTGACCGAGGACAACGCTGTAACCAACAACTTCTACTGTTCCTCCAAACAACTC	3000
Db	2959	CAGCTGACCGAGGACAACGCTGTAACCAACAACTTCTACTGTTCCTCCAAACAACTC	3018
QY	3001	GATGAGGCTTCTGGCGCCAAACGACGATTTGTACACTGCGAAGTGAAGTGGACCACTC	3060
Db	3019	GATGAGGCTTCTGGCGCCAAACGACGATTTGTACACTGCGAAGTGAAGTGGACCACTC	3078
QY	3061	CGCCGGAGATCAGGGAACGAGAGATCGACTTACAGCCAGAAAGCAACGATGGAGGCT	3120
Db	3079	CGCCGGAGATCAGGGAACGAGAGATCGACTTACAGCCAGAAAGCAACGATGGAGGCT	3138
QY	3121	CTGAAGACCACTGTCACATGCTGGAGGAACAGGTATGGAATTTGGAGGCCCTTAAACGAT	3180
Db	3139	CTGAAGACCACTGTCACATGCTGGAGGAACAGGTATGGAATTTGGAGGCCCTTAAACGAT	3198
QY	3181	GAGCTGTAGAAAAAGACGGCGAGTGGAGGCTGGAGAGCGCTCTCGGGTGAATGAGAAA	3240
Db	3199	GAGCTGTAGAAAAAGACGGCGAGTGGAGGCTGGAGAGCGCTCTCGGGTGAATGAGAAA	3258
QY	3241	TCCAGTTTCAGTGTCCGGTTTCAGAGCTCAGAGAAATGCTGGACACCGAGAAAACAGAGC	3300
Db	3259	TCCAGTTTCAGTGTCCGGTTTCAGAGCTCAGAGAAATGCTGGACACCGAGAAAACAGAGC	3318
QY	3301	AGGCGGAGCCGATCAGCGGATCAACCGAGTCTGCCAGGTGTGGAGCTGGCAGTGAAG	3360

Db	3319	AGGGCGAGAGCGCGATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAG	3379
QY	3361	GAGCAAAAGGCTGAGATTCTCGCTCTGCGAGCAGGCTCTCAAAGAGCAGAAGCTCGAAGGCC	3420
Db	3379	GAGCAAAAGGCTGAGATTCTCGCTCTGCGAGCAGGCTCTCAAAGAGCAGAAGCTCGAAGGCC	3438
QY	3421	GAGAGCCTCTCTGCAAAAGCTCAATGACTCGAGAGAAGCATGCTATGCTTGGAAATGAAT	3480
Db	3439	GAGAGCCTCTCTGCAAAAGCTCAATGACTCGAGAGAAGCATGCTATGCTTGGAAATGAAT	3498
QY	3481	GCCCGAAGCTTACAGCGAGAGCTGGAGACTGACGAGAGCTCAAAAGAGGCTTCTCGAA	3540
Db	3499	GCCCGAAGCTTACAGCGAGAGCTGGAGACTGACGAGAGCTCAAAAGAGGCTTCTCGAA	3558
QY	3541	GAGCAAGCAAAATTACAGCAGCAGATGGAGCTGCAGAAAAATCACATTTTCCGCTCTGACT	3600
Db	3559	GAGCAAGCAAAATTACAGCAGCAGATGGAGCTGCAGAAAAATCACATTTTCCGCTCTGACT	3618
QY	3601	CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGTGAATGACTTG	3660
Db	3619	CAAGGACTGCAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGTGAATGACTTG	3678
QY	3661	GAGTATCAGCTGGAAAAACATTACAGTTCTCTATTCTCATGAAAGGTGAAAAATGGAAGGC	3720
Db	3679	GAGTATCAGCTGGAAAAACATTACAGTTCTCTATTCTCATGAAAGGTGAAAAATGGAAGGC	3738
QY	3721	ACTATTCTCAACAAACAAACTCATTTGATTTTCTGCAAGCCAAAAATGACCAACTGCT	3780
Db	3739	ACTATTCTCAACAAACAAACTCATTTGATTTTCTGCAAGCCAAAAATGACCAACTGCT	3798
QY	3781	AAAAAGAAAAAGGTTTATTTAGTCGACGAAAGAGACCTCTGTTTACCCACACAGGTT	3840
Db	3799	AAAAAGAAAAAGGTTTATTTAGTCGACGAAAGAGACCTCTGTTTACCCACACAGGTT	3858
QY	3841	CCTCTGCAGTACAATGAGCTGAAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAG	3900
Db	3859	CCTCTGCAGTACAATGAGCTGAAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAG	3918
QY	3901	CTAGAGAGGCCCTTCAGAGAGCCCGCATCGAGCTCCGGTCCGCCCGGAGGAAAGCTGCC	3960
Db	3919	CTAGAGAGGCCCTTCAGAGAGCCCGCATCGAGCTCCGGTCCGCCCGGAGGAAAGCTGCC	3978
QY	3961	CACGCAAAACAAACGACACCCACACCCATCCACGCCAGCCACCGCAGGACGAGATC	4020
Db	3979	CACGCAAAACAAACGACACCCACACCCATCCACGCCAGCCACCGCAGGACGAGATC	4038
QY	4021	GCCATGTCCGCCATCGTGGGTCCGACAGACACCCAGCCAGCTGCCATGAGCTCTGCGCC	4080
Db	4039	GCCATGTCCGCCATCGTGGGTCCGACAGACACCCAGCCAGCTGCCATGAGCTCTGCGCC	4098
QY	4081	CCGCCATCCAGCCGCGAAGAGAGTCTTCAACTCCAGAGGAATTTAGTCGGGGTCTTAAG	4140
Db	4099	CCGCCATCCAGCCGCGAAGAGAGTCTTCAACTCCAGAGGAATTTAGTCGGGGTCTTAAG	4158
QY	4141	GAACGATGACACACAATATTCCTCACCGATTCAACGTTAGGACTGAAATGCGAGCCACA	4200
Db	4159	GAACGATGACACACAATATTCCTCACCGATTCAACGTTAGGACTGAAATGCGAGCCACA	4218
QY	4201	AAGTGTCTGTGTCTGTGATACCGTGCATTTTGGACGCCAGGCATCCAAATGTCTCGAA	4260
Db	4219	AAGTGTCTGTGTCTGTGATACCGTGCATTTTGGACGCCAGGCATCCAAATGTCTCGAA	4278
QY	4261	TGTGAGGTGATGTGTAACCCCAAGTGCTCCAGTGCTTGTGCCAGCACTCTGGGGTGTGCT	4320
Db	4279	TGTGAGGTGATGTGTAACCCCAAGTGCTCCAGTGCTTGTGCCAGCACTCTGGGGTGTGCT	4338
QY	4321	GCTGAATATGCCACACATTCACCGAGGCTTCTGCGGTGACAAATGAATCTCCCAAGT	4380
Db	4339	GCTGAATATGCCACACATTCACCGAGGCTTCTGCGGTGACAAATGAATCTCCCAAGT	4398
QY	4381	CTCAGACCAAGAGGCCCGACGACGCTTGCACCTTGGAAAGGTGGATGAAGTGTCCCAAG	4440
Db	4399	CTCAGACCAAGAGGCCCGACGACGCTTGCACCTTGGAAAGGTGGATGAAGTGTCCCAAG	4458

QY 4441 AATAACAAACGAGGACACGAAAGGCTGGGACAGGAAGTACATGTGCTCGAGGGATCAAAA 4500  
Db 4459 AATAACAAACGAGGACACGAAAGGCTGGGACAGGAAGTACATGTGCTCGAGGGATCAAAA 4518  
QY 4501 GTCCTCATTTATGACAAATGAAGCCACAGAGAGCTGGACAGAGGCGGTGGAAGAAATTTGAG 4560  
Db 4519 GTCCTCATTTATGACAAATGAAGCCACAGAGAGCTGGACAGAGGCGGTGGAAGAAATTTGAG 4578  
QY 4561 CTGTGCTTCGCGACGGGATGTATCTATTATGATGTCGCGTGTGCTTCCGAACTCGCA 4620  
Db 4579 CTGTGCTTCGCGACGGGATGTATCTATTATGATGTCGCGTGTGCTTCCGAACTCGCA 4638  
QY 4621 AATACAGCCAAAGCAGAGAAAAGCAGAGCTGAATGCTAACTGCTTGGAAATCTCCCTGCTG 4680  
Db 4639 AATACAGCCAAAGCAGAGAAAAGCAGAGCTGAATGCTAACTGCTTGGAAATCTCCCTGCTG 4698  
QY 4681 AAACCTGGAAGTGTATGACCGCTAGACATGAATCAACGCTGCCCTTCAGTGACAGGTG 4740  
Db 4699 AAACCTGGAAGTGTATGACCGCTAGACATGAATCAACGCTGCCCTTCAGTGACAGGTG 4758  
QY 4741 GTGTTGTTGGGACCGAGGAGGCTCTACGCCCTGAATGCTTGAATAAATCTCCCTAAC 4800  
Db 4759 GTGTTGTTGGGACCGAGGAGGCTCTACGCCCTGAATGCTTGAATAAATCTCCCTAAC 4818  
QY 4801 CATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATCAAGACCTGGAGAAGCTA 4860  
Db 4819 CATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATCAAGACCTGGAGAAGCTA 4878  
QY 4861 CTCATGATAGCAGGAGAGAGCGGACATGTGTCTTTGTGAGCTGAAGAAAGTGAACAG 4920  
Db 4879 CTCATGATAGCAGGAGAGAGCGGACATGTGTCTTTGTGAGCTGAAGAAAGTGAACAG 4938  
QY 4921 TCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCACCCAAATTTTGAAGCT 4980  
Db 4939 TCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCACCCAAATTTTGAAGCT 4998  
QY 4981 GTCAAGGGCTGCCACTTTTGGGACAGCAAGATTGAGAGCGGCTCTGCAATCTGTGCA 5040  
Db 4999 GTCAAGGGCTGCCACTTTTGGGACAGCAAGATTGAGAGCGGCTCTGCAATCTGTGCA 5058  
QY 5041 GCCATGCCAGCAAGTGTCTATCTCCGCTACAGCAAAACCTCAGCAAAATCTGCAATC 5100  
Db 5059 GCCATGCCAGCAAGTGTCTATCTCCGCTACAGCAAAACCTCAGCAAAATCTGCAATC 5118  
QY 5101 CGGAAAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTCACCAATTTACAGTATC 5160  
Db 5119 CGGAAAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTCACCAATTTACAGTATC 5178  
QY 5161 CTCATTGGAAACCAATAAATCTACGAATCGACATGAAGCAGTACCGCTCAGGAATTC 5220  
Db 5179 CTCATTGGAAACCAATAAATCTACGAATCGACATGAAGCAGTACCGCTCAGGAATTC 5238  
QY 5221 CTGGATAAGAAATGACCAATTCCTTGGCACCCTGCTGTTTGGCGCTCTTCCAAACAGCTTC 5280  
Db 5239 CTGGATAAGAAATGACCAATTCCTTGGCACCCTGCTGTTTGGCGCTCTTCCAAACAGCTTC 5298  
QY 5281 CCTGTCTCAATCTGCAGTGAACACCGCAGGCGACGAGAGAGTACTGTGTGTTTTC 5340  
Db 5299 CCTGTCTCAATCTGCAGTGAACACCGCAGGCGACGAGAGAGTACTGTGTGTTTTC 5358  
QY 5341 CACGAATTTGGAGTGTCTGTTGATCTTACGGAAGACGTAGCGGACACGATCTCAAG 5400  
Db 5359 CACGAATTTGGAGTGTCTGTTGATCTTACGGAAGACGTAGCGGACACGATCTCAAG 5418  
QY 5401 TGGAGTGCCTTACCTTTGGCCTTTGGCTACAGAGAACCTTATCTGTTTGTGACCCACTTC 5460  
Db 5419 TGGAGTGCCTTACCTTTGGCCTTTGGCTACAGAGAACCTTATCTGTTTGTGACCCACTTC 5478  
QY 5461 AACTCACTGGAATATTGAGATTCAGGACAGCTCTCTCAGCAGGAGCCCTGCCCGAGGG 5520  
Db 5479 AACTCACTGGAATATTGAGATTCAGGACAGCTCTCTCAGCAGGAGCCCTGCCCGAGGG 5538

QY 5521 TACTTGACATCCCGAACCCGCGCTTACCTGGGCGCTGCCATTTCTCAGAGCGATTTAC 5580  
Db 5539 TACTTGACATCCCGAACCCGCGCTTACCTGGGCGCTGCCATTTCTCAGAGCGATTTAC 5598  
QY 5581 TTGGGCTCTCATACAGAGTAAATTAAGGGTCAATTTGCTGCAAGGGAACCTCTGTGAAG 5640  
Db 5599 TTGGGCTCTCATACAGAGTAAATTAAGGGTCAATTTGCTGCAAGGGAACCTCTGTGAAG 5658  
QY 5641 GAGTCCGGCACTCAACACACCCGCGCGCTTACCTGCCAGCAGCAGCCCAACAAAGCGA 5700  
Db 5659 GAGTCCGGCACTCAACACACCCGCGCGCTTACCTGCCAGCAGCAGCCCAACAAAGCGA 5718  
QY 5701 GGCCCAACCCACGTCACAAACGAGCAGTACCAACGCGGTGGCCCTCCAGCCAGCGCGCGCC 5760  
Db 5719 GGCCCAACCCACGTCACAAACGAGCAGTACCAACGCGGTGGCCCTCCAGCCAGCGCGCGCC 5778  
QY 5761 GAAGGCCACGACCCGCGAGAGCCAGCAGCAGCCACCCGCTACCGCAGGCGCGGACC 5820  
Db 5779 GAAGGCCACGACCCGCGAGAGCCAGCAGCAGCCACCCGCTACCGCAGGCGCGGACC 5838  
QY 5821 GAGTGGCGAGGACAGTCTCCTGGCGCGCCCTCGAGCAGAGAGTCCCGCGCGCGG 5880  
Db 5839 GAGTGGCGAGGACAGTCTCCTGGCGCGCCCTCGAGCAGAGAGTCCCGCGCGCGG 5898  
QY 5881 ATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGCGTGTGGAAGACAGCAGCAGCGGC 5940  
Db 5899 ATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGCGTGTGGAAGACAGCAGCAGCGGC 5958  
QY 5941 CGGCTGCTCCGGAGCGGTGAGACCCCGCTGTCCAGGTGAACAGGGAAGAGGCGAG 6000  
Db 5959 CGGCTGCTCCGGAGCGGTGAGAGCCCGCTGTCCAGGTGAACAGGGAAGAGGCGAG 6018  
QY 6001 AGTGCTCTCAAGTTTTCAGGTTTACGTTTAACTGTCACTTATGATGATGGAATGGAAGCTG 6060  
Db 6019 AGTGCTCTCAAGTTTTCAGGTTTAACTGTCACTTATGATGATGGAATGGAAGCTG 6078  
QY 6061 GACACCTGCCAGTAACTGTGCTCAGTCTCAGGATCATCCAGTGAATGAGAAATCCGG 6120  
Db 6079 GACACCTGCCAGTAACTGTGCTCAGTCTCAGGATCATCCAGTGAATGAGAAATCCGG 6138  
QY 6121 CAGCAGGTTGAAAAGTGTGTTCTTGAGAACAGATTTATTCG 6159  
Db 6139 CAGCAGGTTGAAAAGTGTGTTCTGAGAAACAGATTTATTCG 6177

RESULT 5  
US-10-415-011-43  
; Sequence 43, Application US/10415011  
; Publication No. US20040053394A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION  
; APPLICANT: GURURAJAN, Rajagopal  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: CHAWLA, Narinder K.  
; APPLICANT: ELLIOTT, Vicki S.  
; APPLICANT: XU, Yuming  
; APPLICANT: ARVIZU, Chandra S.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: RAMKUMAR, Jayalaxmi  
; APPLICANT: DING, Li  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: HAFALIA, April J.A.  
; APPLICANT: NGUYEN, Dannie B.  
; APPLICANT: GANDHI, Ameena R.  
; APPLICANT: LU, Yan  
; APPLICANT: YUE, Henry  
; APPLICANT: BURFORD, Neil  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: LAL, Preeti G.  
; APPLICANT: RECIPON, Shirley A.  
; APPLICANT: LU, Dying Alina M.  
; APPLICANT: BOROWSKI, Mark L.

```
; APPLICANT: THORNTON, Michael B.
; APPLICANT: SWARNAKER Anita
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0262 USN
; CURRENT APPLICATION NUMBER: US/10/415,011
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/47728
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/242,410
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/244,068
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/245,708
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/247,672
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,565
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,730
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,807
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 43
; LENGTH: 6298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040053394A1 7484498CB1
US-10-415-011-43

Query Match      92.0%; Score 5666.4; DB 13; Length 6298;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 5961; Conservative 0; Mismatches 36; Indels 243; Gaps 3;

QY      1  ATGTGTAAGTTCAAAATATGGAGCGCGGAATCCTTTGGATGCTGGTCTGCTCAACCCATT 60
DB      55  ATGTTGAAGTTCAAAATATGGAGCGCGGAATCCTTTGGATGCTGGTCTGCTCAACCCATT 114
QY      61  GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCCACCTTTATGACTCAA 120
DB      115  GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCCACCTTTATGACTCAA 174
QY      121  CAGCAGATGCTCTCTCTTCCCGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
DB      175  CAGCAGATGCTCTCTCTTCCCGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 234
QY      181  GAATGCACTGAGCTCTCTGATGAAGATTAAGCAGCTGAGCAAATTTGTCCGGAAGTAT 240
DB      235  GAATGCACTGAGCTCTCTGATGAAGATTAAGCAGCTGAGCAAATTTGTCCGGAAGTAT 294
QY      241  TCGGACACATAGCTAGTTACAGAGCTCCAGCCTTCGGGAAAGGACTTGAAGTCTAGA 300
DB      295  TCGGACACCATAGCTAGTTACAGAGCTCCAGCCTTCGGGAAAGGACTTGAAGTCTAGA 354
QY      301  AGTCCTGTAGTGTGCTCACCTTTGCTGAAGTGCAGGTGGTGAAGAGAGAAAGCAACCGG 360
DB      355  AGTCCTGTAGTGTGCTCACCTTTGCTGAAGTGCAGGTGGTGAAGAGAGAAAGCAACCGG 414
QY      361  GACATCTATGCTATGAAGTGAAGAGAGGCTTTATTGGCCCAAGAGAGCTTTTCA 420
DB      415  GACATCTATGCTATGAAGTGAAGAGAGGCTTTATTGGCCCAAGAGAGCTTTTCA 474
QY      421  TTTTGGAGAGAGCGGAAACATATTATCTCGAAGCACAAGCCGCTGGATCCCCCAATTA 480
DB      475  TTTTGGAGAGAGCGGAAACATATTATCTCGAAGCACAAGCCGCTGGATCCCCCAATTA 534
QY      481  CAGTATGCCCTTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGG 540

Db      535  CAGTATGCCCTTTTCAGGACAAAAATCACCTTTATCTGTCATGGAATATCAGCCTGGAGGG 594
QY      541  GACTTGTCTGTCACATTTTGAATAGATATGAGGACAGATTAGATGAAAAACCTGATACAGTTT 600
DB      595  GACTTGTCTGTCACATTTTGAATAGATATGAGGACAGATTAGATGAAAAACCTGATACAGTTT 654
QY      601  TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCCTTCATCTGATGGGATACGTCATCGA 660
DB      655  TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGCCTTCATCTGATGGGATACGTCATCGA 714
QY      661  GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGACACATCAAGCTGGTGGATTTT 720
DB      715  GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGACACATCAAGCTGGTGGATTTT 774
QY      721  GGATCTGCCGCGAAAAATGAATTTCAAAACAAGATGTGAATGCCAAACTCCCGATTGGGACC 780
DB      775  GGATCTGCCGCGAAAAATGAATTTCAAAACAAGATGTGAATGCCAAACTCCCGATTGGGACC 834
QY      781  CCAGATTACATGGCTCCTGGAAGTCTGACTGATGAACGGGATGAAAAGGCACCTAC 840
DB      835  CCAGATTACATGGCTCCTGGAAGTCTGACTGATGAACGGGATGAAAAGGCACCTAC 894
QY      841  GGCCTGGAATGTGACTGGTGGTCAAGTGGGCTGATTGCTCTATGAGATGATTTATGGAGA 900
DB      895  GGCCTGGAATGTGACTGGTGGTCAAGTGGGCTGATTGCTCTATGAGATGATTTATGGAGA 954
QY      901  TCCCCCTTCGAGAGGAACTCTGCCAGAACCTTCAATAACATTAATGATTTCCAGCG 960
DB      955  TCCCCCTTCGAGAGGAACTCTGCCAGAACCTTCAATAACATTAATGATTTCCAGCG 1014
QY      961  TTTTGGAAATTTCCAGATGACCCCAAGATGAGAGTGAATTTCTGATCTGATTTCAAAGC 1020
DB      1015  TTTTGGAAATTTCCAGATGACCCCAAGATGAGAGTGAATTTCTGATCTGATTTCAAAGC 1074
QY      1021  TTGTTGTGGCGCAGAAAAGAGAGACTGAAAGTTTGAAGGCTTTTGTGCGCATCTTTCTTC 1080
DB      1075  TTGTTGTGGCGCAGAAAAGAGAGACTGAAAGTTTGAAGGCTTTTGTGCGCATCTTTCTTC 1134
QY      1081  TCTAAAAATTCAGCTGGAACAACATTCGTAACCTCTCCCTCCCTTCGTTCCACCTCAAG 1140
DB      1135  TCTAAAAATTCAGCTGGAACAACATTCGTAACCTCTCTCCCTCCCTTCGTTCCACCTCAAG 1194
QY      1141  TCTGACGATGACACCTCCAAATTTTGATGAACAGAGAGAAATTCGTGGGTTTCATCTCT 1200
DB      1195  TCTGACGATGACACCTCCAAATTTTGATGAACAGAGAGAAATTCGTGGGTTTCATCTCT 1254
QY      1201  CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTGCGGTTTGTGGGTTTTCG 1260
DB      1255  CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACTGCGGTTTGTGGGTTTTCG 1314
QY      1261  TACAGCAAGCACTGGGGATTCTTGGTAGATCTGAGTCTGTTGTGCGGCTCTGAGCTCC 1320
DB      1315  TACAGCAAGCACTGGGGATTCTTGGTAGATCTGAGTCTGTTGTGCGGCTCTGAGCTCC 1374
QY      1321  CTTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1380
DB      1375  CTTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1434
QY      1381  TCTCAGGACAAGTGTCACAAGATGGAGAGGAAATACCCGGTTTACATCGGAGAGTGTCA 1440
DB      1435  TCTCAGGACAAGTGTCACAAGATGGAGAGGAAATACCCGGTTTACATCGGAGAGTGTCA 1494
QY      1441  GAGTGGAGCTGTGCTTAGTCTAGAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1500
DB      1495  GAGTGGAGCTGTGCTTAGTCTAGAGAGGAGTGGAGCTGAAGCCCTCTGAGACTCAGAGA 1554
QY      1501  TCCTCTCTGAGCAGGACCTTGTCTACTCATCACAGATGACAGTACGTTAAAGCAAGT 1560
DB      1555  TCCTCTCTGAGCAGGACCTTGTCTACTCATCACAGATGACAGTACGTTAAAGCAAGT 1614
QY      1561  TTGGAGCAGACCGATGGAGGTGTCCTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620
```

Db 1615 TTGGAGCAAGCAGCGATGGAGGTGTCCAGGAGGATGCAAAAGCACACTGCTCTCCAT 1674  
Qy 1621 GATATCAGAGCAGACCGGAGCTCCAGAAATCBAAGAGCAGGAGTACCAAGCTCAA 1680  
Db 1675 GATATCAGAGCAGACCGGAGCTCCAGAAATCBAAGAGCAGGAGTACCAAGCTCAA 1734  
Qy 1681 GTGGAAGAAATGAGGTGATGATGAATCAGTTGGAGAGGATCTTGCTTCAGCAAGAGA 1740  
Db 1735 GTGGAAGAAATGAGGTGATGATGAATCAGTTGGAGAGGATCTTGCTTCAGCAAGAGA 1794  
Qy 1741 CGGAGTATCTCTAGCAATCTGAGCTGAGAGAGTCTCGGTTGCTGCTGAAGAAATCAAG 1800  
Db 1795 CGGAGTATCTCTAGCAATCTGAGCTGAGAGAGTCTCGGTTGCTGCTGAAGAAATCAAG 1854  
Qy 1801 CGCAAGCGCACAGATGTCAGATAACTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAA 1860  
Db 1855 CGCAAGCGCACAGATGTCAGATAACTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAA 1914  
Qy 1861 GTGGAGAAATATGCGAAACTGGAGAGATCAATGTGAGCAGCAGCTCAAAATTCAGGAG 1920  
Db 1915 GTGGAGAAATATGCGAAACTGGAGAGATCAATGTGAGCAGCAGCTCAAAATTCAGGAG 1974  
Qy 1921 CTCGAAGGAACTCGAAGAGCTG----- 1945  
Db 1975 CTCGAAGGAACTCGAAGAGCTGTAAAGCCAGCAGGAGGCCACGAGCTGCTGCAG 2034  
Qy 1946 -----CAAAGAGCGAGCCGAGAGGGAGCTGGAGAGCTCGAAGCTCGAAGCCGAGAG 1992  
Db 2035 AATATCCGCGCAGGCAAAAGGAGCGAGCCGAGAGGGAGCTGGAGAGCTCGAAGCCGAGAG 2094  
Qy 1993 GATTCTTCTGAAGGATCAGAAAGAGCTGTGGAGCTGAGGAACCGCGCATCTCTG 2052  
Db 2095 GATTCTTCTGAAGGATCAGAAAGAGCTGTGGAGCTGAGGAACCGCGCATCTCTG 2154  
Qy 2053 GAGAAACAAGGTAAAGAGACTAGAGACCAATGGAGCGCTAGAGAAAACAGACTGAAGGATGAC 2112  
Db 2155 GAGAACAGGTAAAGAGACTAGAGACCAATGGAGCGTAGAGAAAACAGACTGAAGGATGAC 2214  
Qy 2113 ATCCAGACAAAATCCAAACAGATCCAGCAGATGGCTGATATAAATCTCGAGCTCGAAGAG 2172  
Db 2215 ATCCAGACAAAATCCAAACAGATCCAGCAGATGGCTGATATAAATCTCGAGCTCGAAGAG 2274  
Qy 2173 AAACATCGGGAGGCCCAAGCTCTCAGCCAGCAGCTAGAGTGCACTGAAACAGAAAGAG 2232  
Db 2275 AAACATCGGGAGGCCCAAGCTCTCAGCCAGCAGCTAGAGTGCACTGAAACAGAAAGAG 2334  
Qy 2233 CAGCACTATGAGAAAGATTAAGTGTGGACAATCAGATAAAGAAAGACCTGCTGAC 2292  
Db 2335 CAGCACTATGAGAAAGATTAAGTGTGGACAATCAGATAAAGAAAGACCTGCTGAC 2394  
Qy 2293 AAGGACACTGGAGAACATGATGCAGACACGAGAGGAGGCCCATGAGAGGCCAAA 2352  
Db 2395 AAGGACACTGGAGAACATGATGCAGACACGAGAGGAGGCCCATGAGAGGCCAAA 2454  
Qy 2353 ATTCTCAGCGAAACAGAAAGCGATGATCAATGTATGATTCCAAAGATCAGATCCCTGGAA 2412  
Db 2455 ATTCTCAGCGAACAGAAAGCGATGATCAATGTATGATTCCAAAGATCAGATCCCTGGAA 2514  
Qy 2413 CAGAGATTGTGAACTGTCTGAGCCCAATAAATTGACGAAAATAGCAGTCTTTTACC 2472  
Db 2515 CAGAGATTGTGAACTGTCTGAGCCCAATAAATTGACGAAAATAGCAGTCTTTTACC 2574  
Qy 2473 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAAATTTTAC 2532  
Db 2575 CAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCCAACAGAAAATTTTAC 2634  
Qy 2533 CTGGAGACACAGGCTGGAGATTGGAGGCCCAAGAACCGAAAACTGGAGAGCAGCTGGAG 2592  
Db 2635 CTGGAGACACAGGCTGGAGATTGGAGGCCCAAGAACCGAAAACTGGAGAGCAGCTGGAG 2694  
Qy 2593 AAGATCAGCCACCAAGACCACTGACAGATCGGCTGCTGGAATCGAGTGGAGCAAGATTG 2652  
Db 2695 AAGATCAGCCACCAAGACCACTGACAGATCGGCTGCTGGAATCGAGTGGAGCAAGATTG 2754

Qy 2653 CGGAGGTCTAGAGCAGGAGGAGCAAACTGGAGCTCAAGGCCAGCTCAGCAGAG 2712  
Db 2755 CGGAGGTCTAGTAGACACGAGGAGGAGAAACTGGAGCTCAAGGCCAGCTCAGCAGAG 2814  
Qy 2713 CTACAGCTCTCCCTGAGAGCGCGAGTCACTGTTACAGCCCTGACAGCTGACGGCG 2772  
Db 2815 CTACAGCTCTCCCTGAGGAGCGCGAGTCACTGTTACAGCCCTGACAGCTGACGGCG 2874  
Qy 2773 GCCTTGAGAGCCAGCTTCCGCGAGGCGAAGACAGAGCTGGAAGAGACCAAGAGCT 2832  
Db 2875 GCCTTGAGAGCCAGCTTCCGCGAGGCGAAGACAGAGCTGGAAGAGACCAAGAGCT 2934  
Qy 2833 GAAAGAGAGATCAGGCACTCAAGGCAATAGAGATGAATCCAGCGCAAAATTTGATGCT 2892  
Db 2935 GAAAGAGAGATCAGGCACTCAAGGCAATAGAGATGAATCCAGCGCAAAATTTGATGCT 2994  
Qy 2893 CTTTCGTAACAGCTGTACTGTAATCAGACCTCGAGGAGCAGCTAAACAGCTGACCGAG 2952  
Db 2995 CTTTCGTAACAGCTGTACTGTAATCAGACCTCGAGGAGCAGCTAAACAGCTGACCGAG 3054  
Qy 2953 GACAACTGAACTCAACAAACCAAACTTCTTCTTCTCCAAACAACTCGATGAGGTTCT 3012  
Db 3055 GACAACTGAACTCAACAAACCAAACTTCTTCTTCTCCAAACAACTCGATGAGGTTCT 3114  
Qy 3013 GCGCCAAACAGAGAGTTGTAACTCGAAGTGAAGTGAAGCTCTCCGCGGGAGATC 3072  
Db 3115 GCGCCAAACAGAGAGTTGTAACTCGAAGTGAAGTGAAGCTCTCCGCGGGAGATC 3174  
Qy 3073 ACGGAACGAGAGATGACGTTACAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG 3132  
Db 3175 ACGGAACGAGAGATGACGTTACAGCCAGAAACAAACGATGGAGGCTCTGAAGACCAAG 3234  
Qy 3133 TGCACCAATGTGAGAGAAACAGGTCATGATTTGGAGCCCTAAACGATGAGCTGCTAGAA 3192  
Db 3235 TGCACCAATGTGAGAGAAACAGGTCATGATTTGGAGCCCTAAACGATGAGCTGCTAGAA 3294  
Qy 3193 AAAGAGCGGAGTGGAGGCTCGAGGAGCTCTCGGTGATGAGAATCCAGTTTGGAG 3252  
Db 3295 AAAGAGCGGAGTGGAGGCTCGAGGAGCTCTCGGTGATGAGAATCCAGTTTGGAG 3354  
Qy 3253 TGTCCGTTTCGAGAGCTGACAGAAATGCTGGACAACGAGAAAACAGAGCGGCGAGAGCC 3312  
Db 3355 TGTCCGTTTCGAGAGCTGACAGAAATGCTGGACAACGAGAAAACAGAGCGGCGAGAGCC 3414  
Qy 3313 GATCAGCGGATCAGGAGTCTCCGAGTGGTGGAGCTGAGTGGAGGAGCACAAGGCT 3372  
Db 3415 GATCAGCGGATCAGGAGTCTCCGAGTGGTGGAGCTGAGTGGAGGAGCACAAGGCT 3474  
Qy 3373 GAGATTCTCGCTCTGACAGAGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3432  
Db 3475 GAGATTCTCGCTCTGACAGAGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3534  
Qy 3433 GACAACTCAATGACCTGGAGAAAGCATGCTTATGCTTGAATGAATGACCGAGCTTA 3492  
Db 3535 GACAAGCTCAATGACCTGGAGAAAGCATGCTTATGCTTGAATGAATGACCGAGCTTA 3594  
Qy 3493 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAACAGAGGCTTCTGGAAGAGCAGCCAAA 3552  
Db 3595 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAACAGAGGCTTCTGGAAGAGCAGCCAAA 3654  
Qy 3553 TTACAGCAGCAGATGGACCTGCAGAAAATCACAATTTCCGCTCTGACTCAAGGACTGAA 3612  
Db 3655 TTACAGCAGCAGATGGACCTGCAGAAAATCACAATTTCCGCTCTGACTCAAGGACTGAA 3714  
Qy 3613 GAACTCTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTGAAGTGGAGTATCAGCTG 3672  
Db 3715 GAACTCTAGATCGGCTGATCTTACTGAAGACAGAAAGAGTGAAGTGGAGTATCAGCTG 3774  
Qy 3673 GAAAACATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGCACCTATTCTCAA 3732  
Db 3775 GAAAACATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAATGGAAGGCACCTATTCTCAA 3834

Qy 3733 CAAACCAACTCATGTTTCTGCAAGCCAAATGAGCAACCTGCTCTAAAAGAAAAAG 3792  
Db 3835 CAAACCAACTCATGTTTCTGCAAGCCAAATGAGCAACCTGCTCTAAAAGAAAAAG 3892  
Qy 3793 GGTATTATTTAGTCGACGGAAAGAGGACCTGCTTTACCCACACAGTTCTCTGCAAGTAC 3852  
Db 3893 -----AGGTTCTCTGCAAGTAC 3909  
Qy 3853 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTTAGAGGAAGCC 3912  
Db 3910 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTCAGAGCTTAGAGGAAGCC 3969  
Qy 3913 CTTTCAGAAAGACCGCATCGAGCTCCGCTCCGCCCGGAGGAAGCTGCCACCGCAAGCA 3972  
Db 3970 CTTTCAGAAAGACCGCATCGAGCTCCGCTCCGCCCGGAGGAAGCTGCCACCGCAAGCA 4029  
Qy 3973 ACGACCAACCAACACCATCCAGCCAGCCACCGAGGAGGAGATCGCCATGTCGCCG 4032  
Db 4030 ACGACCAACCAACACCATCCAGCCAGCCACCGAGGAGGAGATCGCCATGTCGCCG 4089  
Qy 4033 ATCTGCGGTGCGCCAGAGCAACAGCCAGTGCATGAGCTGCTGCCCCCGCATCCAGC 4092  
Db 4090 ATCTGCGGTGCGCCAGAGCAACAGCCAGTGCATGAGCTGCTGCCCCCGCATCCAGC 4149  
Qy 4093 CGCAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCCGCGCTTTAAGGAAAGCATGAC 4152  
Db 4150 CGCAGAAAGAGTCTTCACTCCAGAGGAATTTAGTCCGCGCTTTAAGGAAAGCATGAC 4209  
Qy 4153 CACAATATTTCTCACCGATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTCTGTG 4212  
Db 4210 CACAATATTTCTCACCGATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTCTGTG 4269  
Qy 4213 TGTCTGGATACCGTGCACATTGAGCCAGGAGATCCAAATGCTCGAATGTGAGTGATG 4272  
Db 4270 TGTCTGGATACCGTGCACATTGAGCCAGGAGATCCAAATGCTCGAATGTGAGTGATG 4329  
Qy 4273 TGTCAACCCCAAGTGTCTCACGCTCTTCCAGCCACCTCGGGTGTGCTCTCAATATGCC 4332  
Db 4330 TGTCAACCCCAAGTGTCTCACGCTCTTCCAGCCACCTCGGGTGTGCTCTCAATATGCC 4389  
Qy 4333 ACACATTCACCGAGGCTTGTCCGTGACAAATGAATCCCAAGTTCAGACCAAG 4392  
Db 4390 ACACATTCACCGAGGCTTGTCCGTGACAAATGAATCCCAAGTTCAGACCAAG 4449  
Qy 4393 GAGCCAGCAGAGCTTGACCTTGGAGGAGTGGATGAAGTGCAGGAATTAACAAACGA 4452  
Db 4450 GAGCCAGCAGAGCTTGACCTTGGAGGAGTGGATGAAGTGCAGGAATTAACAAACGA 4509  
Qy 4453 GGACAGCAAGGCTGGACAGGAAGTACATTGTCTGGAGGATCAAAAGTCTCTATTAT 4512  
Db 4510 GGACAGCAAGGCTGGACAGGAAGTACATTGTCTGGAGGATCAAAAGTCTCTATTAT 4569  
Qy 4513 GACAAATGAAGCCAGAGAACTGGACAGAGCCGGTGGAGAAATTTGAGCTGTCCTTCCC 4572  
Db 4570 GACAAATGAAGCCAGAGAACTGGACAGAGCCGGTGGAGAAATTTGAGCTGTCCTTCCC 4629  
Qy 4573 GAGCGGATGTATCTATTATCATGCTGCGTGTGCTTCCGAATCGCAAAATACAGCCAAA 4632  
Db 4630 GAGCGGATGTATCTATTATCATGCTGCGTGTGCTTCCGAATCGCAAAATACAGCCAAA 4689  
Qy 4633 GCA----- 4635  
Db 4690 GCAGATGTCCCATACATACTGAAGATGGAATCTCCCGCACACACCTGCTGGCCCGG 4749  
Qy 4636 ----- 4635  
Db 4750 AGAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCCCTTA 4809  
Qy 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4662  
Db 4810 GAATCAGTTGTCCAGGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAAACTG 4869  
Qy 4663 CTTGGAAACTCCTGCTGTAAGTGGAGGAGTATGACCGTCTAGACATGAACCTGCACGCTG 4722

Db 4870 CTTGAAACTCCTGCTGTAAGTGGAGGATGATGACCGTCTAGACATGAACCTGCACGCTG 4929  
Qy 4723 CCCTTCAGTACGACAGTGGTGTGGTGGGACCCAGGAAGGGCTCTACGCCCTGAATGTC 4782  
Db 4930 CCCTTCAGTACGACAGTGGTGTGGTGGGACCCAGGAAGGGCTCTACGCCCTGAATGTC 4989  
Qy 4783 TTGAAAACTCCCTAAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATTC 4842  
Db 4990 TTGAAAACTCCCTAAACCCATGTCCAGGAATTTGGAGCAGTCTTCCAAATTTATTTATTC 5049  
Qy 4843 AAGACCTTGAAGAGCTTACTCATGATAGCAGGAAGAGCGGCACTGTGTCTTTGTGGAC 4902  
Db 5050 AAGACCTTGAAGAGCTTACTCATGATAGCAGGAAGAGCGGCACTGTGTCTTTGTGGAC 5109  
Qy 4903 GTGAAGAAAGTGAACAGTCCCTGGCCCACTGCCACCTGCTGCCAGCCCGGACATCTCA 4962  
Db 5110 GTGAAGAAAGTGAACAGTCCCTGGCCCACTGCCACCTGCTGCCAGCCCGGACATCTCA 5169  
Qy 4963 CCCAACATTTTGAAGCTGTCAAGGGTGCACCTTGTTCGGGAGGCAAGATTTGAGAAC 5022  
Db 5170 CCCAACATTTTGAAGCTGTCAAGGGTGCACCTTGTTCGGGAGGCAAGATTTGAGAAC 5229  
Qy 5023 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTGTGCTATTCTCCGCTACACGAAAC 5082  
Db 5230 GGGCTCTGCATCTGTGCAGCCATGCCAGCAAAAGTGTGCTATTCTCCGCTACACGAAAC 5289  
Qy 5083 CTCAGCAAACTACTGCATTCGCGAAAGAGATAGACCTCAGAGCCCTGCAGCTGTATCCAC 5142  
Db 5290 CTCAGCAAACTACTGCATTCGCGAAAGAGATAGACCTCAGAGCCCTGCAGCTGTATCCAC 5349  
Qy 5143 TTCAACCAATTACAGTATCTCANTGGNAACCAATAAATTTACCAAAATCGACATGAAGCAG 5202  
Db 5350 TTCAACCAATTACAGTATCTCANTGGNAACCAATAAATTTACCAAAATCGACATGAAGCAG 5409  
Qy 5203 TACACGCTCCAGAGAAATTCCTGGATAGAATGACATTCCTTTGGCAGCTGCTGTGTTGCC 5262  
Db 5410 TACACGCTCCAGAGAAATTCCTGGATAGAATGACATTCCTTTGGCAGCTGCTGTGTTGCC 5469  
Qy 5263 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGAGGCGAGAG 5322  
Db 5470 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTCAGGTGAACAGCGAGGCGAGAG 5529  
Qy 5323 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTGTGTGATTTTACGGAAGACGTAGC 5382  
Db 5530 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTGTGTGATTTTACGGAAGACGTAGC 5589  
Qy 5383 CGCACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTTGCCCTACAGAGAACCTTAT 5442  
Db 5590 CGCACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTTGCCCTACAGAGAACCTTAT 5649  
Qy 5443 CTGTTTGTGACCCACTTCAACTCAGTTCAGTAAATTTAGATCCAGGACGCTCCTCAGCA 5502  
Db 5650 CTGTTTGTGACCCACTTCAACTCAGTTCAGTAAATTTAGATCCAGGACGCTCCTCAGCA 5709  
Qy 5503 GGGACCCCTCCCGAGCGTACCTGGACATCCCGAAACCGCGCTTACCTGGGCGCTTGCATT 5562  
Db 5710 GGGACCCCTCCCGAGCGTACCTGGACATCCCGAAACCGCGCTTACCTGGGCGCTTGCATT 5769  
Qy 5563 TCCTCAGGAGCGATTTACTTGGCGTCTCTATACAGGATAAATTAAGGTCTATTGCTGC 5622  
Db 5770 TCCTCAGGAGCGATTTACTTGGCGTCTCTATACAGGATAAATTAAGGTCTATTGCTGC 5829  
Qy 5623 AAGGAAACCTCTGTGAAGGAGTCCGCACTGAAACAGACACCGGGGCGCTTCCACCTCCCGC 5682  
Db 5830 AAGGAAACCTCTGTGAAGGAGTCCGCACTGAAACAGACACCGGGGCGCTTCCACCTCCCGC 5889  
Qy 5683 AGCAGCCCAAGAGCAGGCGCCACCCAGTACAAAGAGACATACCAAGCGCGTGGCC 5742  
Db 5890 AGCAGCCCAAGAGCAGGCGCCACCCAGTACAAAGAGACATACCAAGCGCGTGGCC 5949  
Qy 5743 TCAGGCGCAGCGCGCCGAAAGGCCAGCCAGCCCGGAGGAGCAAGCACACCCACCGC 5802

Db 5950 TCCAGCCAGGCGCCCGAGGCCAGCCAGCCAGCCAGGAGCCAGCACACCCACCGC 6009  
Qy 5803 TACCCGAGGGCGACGAGTGCAGGAGCAAGTCTCTGCGCGCCCTGGAGCGA 5862  
Db 6010 TACCCGAGGGCGGACGAGTGCAGGAGCAAGTCTCTGCGCGCCCTGGAGCGA 6069  
Qy 5863 GAGAAAGTCCCCCGCGGATGTCTAGCAGCGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5922  
Db 6070 GAGAAAGTCCCCCGCGGATGTCTAGCAGCGCGGAGAGCGGTCCCGCGGAGGCTGTTT 6129  
Qy 5923 GAAGCAGCAGAGGGCGCGGTGCTGCGGGAGCGGTGAGGACCCCGCTGTCCAGGTG 5982  
Db 6130 GAAGCAGCAGAGGGCGCGGTGCTGCGGGAGCGGTGAGGACCCCGCTGTCCAGGTG 6189  
Qy 5983 AACAAAGGAGAGGCGAGAGTCCCTCTCAAGTTTTCAGGTTTAACTGTCACTATTAT 6042  
Db 6190 AACAAAGTCTGGACAGTCTTCAATATTAATCTCAGCCAGAGAAACCAATCTCTCATCT 6249

RESULT 6  
US-10-028-946-1  
; Sequence 1, Application US/10028946  
; Publication No. US20020123622A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Miranda, Maricar  
; APPLICANT: Friddle, Carl Johan  
; TITLE OF INVENTION: No. US20020123622A1el Human Kinases and Polynucleotides Encoding  
; FILE REFERENCE: LEX-0289-USA  
; CURRENT APPLICATION NUMBER: US/10/028,946  
; CURRENT FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: US 60/258,335  
; PRIOR FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 6165  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-10-028-946-1

Query Match 91.9%; Score 5661.4; DB 14; Length 6165;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

Qy 1 ATGTTGAAGTTCAAATATGAGCGGGGAATCCTTTGGATGCTGTGGTCTGCTGAACCCCAT 60  
Db 1 ATGTTGAAGTTCAAATATGAGCGGGGAATCCTTTGGATGCTGTGGTCTGCTGAACCCCAT 60  
Qy 61 GCCAGCGGGCTCCAGGCTGAATCTCTTCCAGGGGAACCCCTTTATGACTCAA 120  
Db 61 GCCAGCGGGCTCCAGGCTGAATCTCTTCCAGGGGAACCCCTTTATGACTCAA 120  
Qy 121 CAGCAGATGCTCTCTCTTTCCGAGAGGGAATATAGATGCCCTCTTTGTTCTTTTCAA 180  
Db 121 CAGCAGATGCTCTCTCTTTCCGAGAGGGAATATAGATGCCCTCTTTGTTCTTTTCAA 180  
Qy 181 GAATGAGTCAAGCTCTCTATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAGTAT 240  
Db 181 GAATGAGTCAAGCTCTCTATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAGTAT 240  
Qy 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCAAGGACTTCGAAAGTCAGA 300  
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTCGGCAAGGACTTCGAAAGTCAGA 300  
Qy 301 AGTCTTGTAGTGTGGTCACTTTGCTGTAAGTGCAGGTGTAAGAGAGAAACCCGGG 360  
Db 301 AGTCTTGTAGTGTGGTCACTTTGCTGTAAGTGCAGGTGTAAGAGAGAAACCCGGG 360  
Qy 361 GACATCTATGCTATGAAGTATGAAGAGAGGCTTTATTTGCCCGAGGAGGTTTCA 420  
Db 361 GACATCTATGCTATGAAGTATGAAGAGAGGCTTTATTTGCCCGAGGAGGTTTCA 420

Qy 421 TTTTGTGAGGAGAGCGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480  
Db 421 TTTTGTGAGGAGAGCGGGAACATATTATCTCGAAGCACAAAGCCCGTGGATCCCCCAATTA 480  
Qy 481 CAGTATGCCCTTTAGAGCAAAAATCACCCTTATCTGATGGAGGAATATCAGCCTGGAGGG 540  
Db 481 CAGTATGCCCTTTAGAGCAAAAATCACCCTTATCTGATGGAGGAATATCAGCCTGGAGGG 540  
Qy 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGTATACAGTTT 600  
Db 541 GACTTGTGTCACCTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGTATACAGTTT 600  
Qy 601 TACCTAGCTCAGCTGATTTTGGCTGTTTCAAGCGTTTCACTGATGGATACGTGTCATCGA 660  
Db 601 TACCTAGCTCAGCTGATTTTGGCTGTTTCAAGCGTTTCACTGATGGATACGTGTCATCGA 660  
Qy 661 GACATCAAGCCTCAGAAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720  
Db 661 GACATCAAGCCTCAGAAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720  
Qy 721 GGATCTCCGCGAAATGAATTCAAACAAGATGGTGAATCCCAACTCCCGATTTGGGACC 780  
Db 721 GGATCTCCGCGAAATGAATTCAAACAAGATGGTGAATCCCAACTCCCGATTTGGGACC 780  
Qy 781 CCAGATTACATGCTCCTGAGTGTCTGACTGTGATGAACGGGATGGAAGGCACCTAC 840  
Db 781 CCAGATTACATGCTCCTGAGTGTCTGACTGTGATGAACGGGATGGAAGGCACCTAC 840  
Qy 841 GGCCTGACTGTGACTGCTGGTGGTGTGATTCCTTATGAGATGATTTATGGGAGA 900  
Db 841 GGCCTGACTGTGACTGCTGGTGGTGTGATTCCTTATGAGATGATTTATGGGAGA 900  
Qy 901 TCCCCCTTCCGAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTAATTAATTCACGGG 960  
Db 901 TCCCCCTTCCGAGAGGGAACCTCTGCCAGAACCTTCAATTAACATTAATTAATTCACGGG 960  
Qy 961 TTTTGTGAATTTCCAGATGATCCCAAGTGAGCAGTCACTTCTTGTATCTGATTCATCAAGC 1020  
Db 961 TTTTGTGAATTTCCAGATGATCCCAAGTGAGCAGTCACTTCTTGTATCTGATTCATCAAGC 1020  
Qy 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTTTGCTGCCATCCCTTCTTC 1080  
Db 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAAGTTTGAAGTCTTTGCTGCCATCCCTTCTTC 1080  
Qy 1081 TCTAAATTTGACTGGAACACATTCGTAACCTCTCTCCCTCCCTTCGTTCCCAACCTCAAG 1140  
Db 1081 TCTAAATTTGACTGGAACACATTCGTAACCTCTCTCTCCCTTCGTTCCCAACCTCAAG 1140  
Qy 1141 TCTCAGCATGACACCTCCAAATTTTGTATGAACCCAGAGAGAAATTCGTGGGTTTCATCTCT 1200  
Db 1141 TCTCAGCATGACACCTCCAAATTTTGTATGAACCCAGAGAGAAATTCGTGGGTTTCATCTCT 1200  
Qy 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTGAGAACTGCCGTTTGTGGGTTTTTCG 1260  
Db 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTGAGAACTGCCGTTTGTGGGTTTTTCG 1260  
Qy 1261 TACAGCAAGGCACGTGGGATTTCTTGGTAGATCTGAGTCTGTTGTGCGGTTCTGAGTCTC 1320  
Db 1261 TACAGCAAGGCACGTGGGATTTCTTGGTAGATCTGAGTCTGTTGTGCGGTTCTGAGTCTC 1320  
Qy 1321 CCTGCCAAGACTAGCTCCATGGAAAAAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380  
Db 1321 CCTGCCAAGACTAGCTCCATGGAAAAAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380  
Qy 1381 TCTCAGGACAAAGTGTCAAGATGGAGCAAAATGACCCGTTTACATCGGAGGTCTCA 1440  
Db 1381 TCTCAGGACAAAGTGTCAAGATGGAGCAAAATGACCCGTTTACATCGGAGGTCTCA 1440  
Qy 1441 GAGGTGGAGGCTGTCTTAGTTCAGAGAGGAGTGGAGTGAAGGCTCTGAGACTCAGAGA 1500  
Db 1441 GAGGTGGAGGCTGTCTTAGTTCAGAGAGGAGTGGAGTGAAGGCTCTGAGACTCAGAGA 1500  
Qy 1501 TCCCTCTCTGGAGGAGGACCTTGTCTACCTACATCAACAATGCAGTAGCTTAAAGCGAGT 1560

1501 TCCTCTCGAGCAGGACCTTGTACTACATCACAGAATGCAGTAGCTTAAAGCGAAGT 1560  
1561 TTGGAGCAAGCAGGATGGAGGTGTCCTCAGCAGGATGACAAAGCACTGCAGCTTCTCCAT 1620  
1561 TTGGAGCAAGCAGGATGGAGGTGTCCTCAGCAGGATGACAAAGCACTGCAGCTTCTCCAT 1620  
1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGGATGACCAAGGCTCAA 1680  
1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGGATGACCAAGGCTCAA 1680  
1681 GTGGAAGAAATGAGGTGATGATGAATCAAGTTCGGAAGAGATCTTCTCAGCAGAGGA 1740  
1681 GTGGAAGAAATGAGGTGATGATGAATCAAGTTCGGAAGAGATCTTCTCAGCAGAGGA 1740  
1741 CGAGTGATCTTACGAATCTGAGCTGAGAGATCTCGGCTTGCTGCTGAAGAATTCAG 1800  
1741 CGAGTGATCTTACGAATCTGAGCTGAGAGATCTCGGCTTGCTGCTGAAGAATTCAG 1800  
1801 CGGAAGCGACAGAAATGTCAGCAATAACTGTTGAAGCTTAAGATCAAGGAGAGCTGAA 1860  
1801 CGGAAGCGACAGAAATGTCAGCAATAACTGTTGAAGCTTAAGATCAAGGAGAGCTGAA 1860  
1861 GTGGGAGAAATATCGGAATCTGGAGAAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
1861 GTGGGAGAAATATCGGAATCTGGAGAAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
1921 CTCAGAGAGAACTGGAGAAAGGT----- 1944  
1921 CTCAGAGAGAACTGGAGAAAGGTGTAAAGCCAGCAGCAGGAGGCCACCCAGAGCTGTGCGAG 1980  
1945 -----CAGAGAGCGAGCGGAGAGGAGCTCGAGAAGCTGCAGAACCGAGAG 1992  
1981 AATATCGGCAGCAAGAGCGAGCGAGAGGAGCTGGAGAGCTGGAGAGCTGCAGAACCGAGAG 2040  
1993 GATCTCTGAAAGCATCAGAAAGAGCTGGTGGAGCTTGAGGAACGCCGCCATCTCTG 2052  
2041 GATCTCTGAAAGCATCAGAAAGAGCTGGTGGAGCTTGAGGAACGCCGCCATCTCTG 2100  
2053 GAGAACAGGTAAAGAGCTAGAGACCATGGAGCGTGAAGAGAAACAGACTGAAGATGAC 2112  
2101 GAGAACAGGTAAAGAGCTAGAGACCATGGAGCGTGAAGAGAAACAGACTGAAGATGAC 2160  
2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGAGCTCGAAGAG 2172  
2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGAGCTCGAAGAG 2220  
2173 AACATCGGAGGCCCAAGTCTCAGCCAGACCTAGAGTGCACTGAAACAGAAAGAG 2232  
2221 AACATCGGAGGCCCAAGTCTCAGCCAGACCTAGAGTGCACTGAAACAGAAAGAG 2280  
2233 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATAAAGAAAGACCTGGCTGAC 2292  
2281 CAGCACTATGAGGAAAGATTAAGTGTGGACAATCAGATAAAGAAAGACCTGGCTGAC 2340  
2293 AAGGAGACCTGGAGAACATGATCGAGACACGAGGAGGAGGCCCATGAGAGGGCAAA 2352  
2341 AAGGAGACCTGGAGAACATGATCGAGACACGAGGAGGAGGCCCATGAGAGGGCAAA 2400  
2353 ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGATGATTCAGAGTCCAGATCCCTGGAA 2412  
2401 ATTCTCAGCGAACAGAGCGGATGATCAATGCTATGATGATTCAGAGTCCAGATCCCTGGAA 2460  
2413 CAGAGGATGTGGAATGCTGAAAGCCAAATAACTTGACGCAAAATAGCAGTCTTTTACC 2472  
2461 CAGAGGATGTGGAATGCTGAAAGCCAAATAACTTGACGCAAAATAGCAGTCTTTTACC 2520  
2473 CAGAGGACATGAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2532  
2521 CAGAGGACATGAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2580  
2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCGAAACTGGAGGAGCAGCTGGAG 2592

2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG 2640  
2593 AAGATCAGCCACCAAGACCACAGTGACAAGAAATCGGCTGCTGGAACTGGAGACAAAGATTG 2652  
2641 AAGATCAGCCACCAAGACCACAGTGACAAGAAATCGGCTGCTGGAACTGGAGACAAAGATTG 2700  
2653 CGGAGGTGATCTGTAGAGCAGCAGGAGCAGAACTGGAGCTCAAGCCGAGCTCAGAG 2712  
2701 CGGAGGTGATCTGTAGAGCAGCAGGAGCAGAACTGGAGCTCAAGCCGAGCTCAGAG 2760  
2713 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCCCTGACAGGCTGACAGGCG 2772  
2761 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCCCTGACAGGCTGACAGGCG 2820  
2773 GCCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCAAGCAGAAAGCT 2832  
2821 GCCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCAAGCAGAAAGCT 2880  
2833 GAAGAGGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAATTTGATGCT 2892  
2881 GAAGAGGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCGCAATTTGATGCT 2940  
2893 CTTCTAACAGCTGTACTGTAAATCACAGACTCGGAGGAGCAGCTAAACAGCTGACCGAG 2952  
2941 CTTCTAACAGCTGTACTGTAAATCACAGACTCGGAGGAGCAGCTAAACAGCTGACCGAG 3000  
2953 GACAACTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3012  
3001 GACAACTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060  
3013 GGCCCAACCAAGAGATGTGTAACCTGCGAAGTGAAGTGAACCATCTCCGCCGGGAGATC 3072  
3061 GGCCCAACCAAGAGATGTGTAACCTGCGAAGTGAAGTGAACCATCTCCGCCGGGAGATC 3120  
3073 ACGAAACAGAGATGAGCTTACAGCCAGAAACAGATGAGGCTCTGAAGACACAG 3132  
3121 ACGAAACAGAGATGAGCTTACAGCCAGAAACAGATGAGGCTCTGAAGACACAG 3180  
3133 TGCACTGCTGGAGGAAACAGGCTCATGATTTGAGGCCCTAAACATGAGCTGTAGAA 3192  
3181 TGCACTGCTGGAGGAAACAGGCTCATGATTTGAGGCCCTAAACATGAGCTGTAGAA 3240  
3193 AAAGAGCGCAGTGGGAGGCTCGAGAGCGTCTCGGTGATGAGAAATCCCAAGTTTGA 3252  
3241 AAAGAGCGCAGTGGGAGGCTCGAGAGCGTCTCGGTGATGAGAAATCCCAAGTTTGA 3300  
3253 TGTCCGTTGAGAGCTGACAGAAATGCTGGACACCGAAGAACAGAGCAGGCGAGAGCC 3312  
3301 TGTCCGTTGAGAGCTGACAGAAATGCTGGACACCGAAGAACAGAGCAGGCGAGAGCC 3360  
3313 GATCAGCGGATCACCGAGTCTCCGAGGTGGTGGAGCTGCGAGTGAAGGAGCAAGGCT 3372  
3361 GATCAGCGGATCACCGAGTCTCCGAGGTGGTGGAGCTGCGAGTGAAGGAGCAAGGCT 3420  
3373 GAGATTCCTGCTGACAGAGGCTCTCAAAGACAGAGCTGAAGCCCGAGGAGCTCTCT 3432  
3421 GAGATTCCTGCTGACAGAGGCTCTCAAAGACAGAGCTGAAGCCCGAGGAGCTCTCT 3480  
3433 GACAAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCCAAGCTTA 3492  
3481 GACAAGCTCAATGACCTGGAGAAAGAGCATGCTATGCTTGAATGAATGCCCAAGCTTA 3540  
3493 CAGCAGAGCTGAGACTGAACAGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAAGCCAA 3552  
3541 CAGCAGAGCTGAGACTGAACAGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600  
3553 TTACAGCAGAGAGTGGACCTGACAGAAATACATTTTCCGCTGACTCAAGGAGCTCAA 3612  
3601 TTACAGCAGAGAGTGGACCTGACAGAAATACATTTTCCGCTGACTCAAGGAGCTCAA 3660  
3613 GAAAGCTTAGATCGGCTGATCTTCTGAAGACAGAAAGAGTGAAGTGAAGTATCAGCTG 3672  
3661 GAAAGCTTAGATCGGCTGATCTTCTGAAGACAGAAAGAGTGAAGTGAAGTATCAGCTG 3720

QY 3673 GAAACATTCAGTCTCTATTTCTCATGAAAGGTGAAAATGGAAGGCACATTTTCTCAA 3732  
Db 3721 GAAACATTCAGTCTCTATTTCTCATGAAAGGTGAAAATGGAAGGCACATTTTCTCAA 3780  
QY 3733 GAAACCAAATCATTTGATTTCTGCAAGCCAAATGGAACCACTGCTTAAAAAGAAAAG 3792  
Db 3781 CAAACCAAATCATTTGATTTCTGCAAGCCAAATGGAACCACTGCTTAAAAAGAAAAG 3838  
QY 3793 GGTTTATTTAGTCGACGGAAGAGGACCCTGCTTTTACCACACACAGGTTTCCTCTGCAATAC 3852  
Db 3839 -----AGTTTCTCTGCAATAC 3855  
QY 3853 AATGAGCTGAAGTCCGCTCGAGAAAGAGAAAGCTCGCTGTGCAAGCTAGAGAAAGCC 3912  
Db 3856 AATGAGCTGAAGTCCGCTCGAGAAAGAGAAAGCTCGCTGTGCAAGCTAGAGAAAGCC 3915  
QY 3913 CTTTCAAGACCCGATCGAGCTCCGCTCCGCGGAGGAGAGCTGCCACCGCAAGCA 3972  
Db 3916 CTTTCAAGACCCGATCGAGCTCCGCTCCGCGGAGGAGAGCTGCCACCGCAAGCA 3975  
QY 3973 ACGGACCAACCCATCCACGCCAGCCACCGCGAGGCAAGATCGCCATGTCGCC 4032  
Db 3976 ACGGACCAACCCATCCACGCCAGCCACCGCGAGGCAAGATCGCCATGTCGCC 4035  
QY 4033 ATCTGCGGTCCGAGAGCAGCCAGTCCAGTSCATGAGCTGCTGCGCCGCGCATCCAGC 4092  
Db 4036 ATCTGCGGTCCGAGAGCAGCCAGTCCAGTSCATGAGCTGCTGCGCCGCGCATCCAGC 4095  
QY 4093 CCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCCGCGCTTTAAGGAAACGATGCAC 4152  
Db 4096 CCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCCGCGCTTTAAGGAAACGATGCAC 4155  
QY 4153 CACAATATTCCTCAGGATTTCACTGAGTGAACATCGAGGCCACAAAGTGTGCTGTG 4212  
Db 4156 CACAATATTCCTCAGGATTTCACTGAGTGAACATCGAGGCCACAAAGTGTGCTGTG 4215  
QY 4213 TGTCTGGATACCGTGCACCTTGGAGCCAGGCATCCAAATGTCGAAATGTCAGTGTATG 4272  
Db 4216 TGTCTGGATACCGTGCACCTTGGAGCCAGGCATCCAAATGTCGAAATGTCAGTGTATG 4275  
QY 4273 TGTACCCCAAGTGTCCACGTGTGCGAGCCACTCGGCTTCCCTGCTGAATATGCC 4332  
Db 4276 TGTACCCCAAGTGTCCACGTGTGCGAGCCACTCGGCTTCCCTGCTGAATATGCC 4335  
QY 4333 ACACATTCACCGAGCCTTCTCCGTGACAAATGAACTCCCGAGTCTCCAGACCAG 4392  
Db 4336 ACACATTCACCGAGCCTTCTCCGTGACAAATGAACTCCCGAGTCTCCAGACCAG 4395  
QY 4393 GAGCCAGCAGCAGCTTGCACTGGAAGGCTGGATGAAGGTGCCAGGAATAACAAACGA 4452  
Db 4396 GAGCCAGCAGCAGCTTGCACTGGAAGGCTGGATGAAGGTGCCAGGAATAACAAACGA 4455  
QY 4453 GGACAGCAAGGCTGGACAGGAAGTACATTTCTCTGGAGGATCAAAAGTCTCATTTAT 4512  
Db 4456 GGACAGCAAGGCTGGACAGGAAGTACATTTCTCTGGAGGATCAAAAGTCTCATTTAT 4515  
QY 4513 GACATGAAGCCAGAGAAGCTGACAGAGCCGCTGGAAGAAATTTGAGCTGTCCTTCCC 4572  
Db 4516 GACATGAAGCCAGAGAAGCTGACAGAGCCGCTGGAAGAAATTTGAGCTGTCCTTCCC 4575  
QY 4573 GACGGGATGTATCTATTATGATGCTGCTGCTTCCGAACTCGCAATACAGCCAAA 4632  
Db 4576 GACGGGATGTATCTATTATGATGCTGCTTCCGAACTCGCAATACAGCCAAA 4635  
QY 4633 GCA----- 4635  
Db 4636 GCAGATGTCCCATACATCTGAAGATGGAAATCTCACCCGCACACCACTGTGGCCCGGG 4695  
QY 4636 ----- 4635  
Db 4696 AGAACCTCTACTGTAGTCCACGCTTCCCTTGACAAACAGCGGTGGGTACCCGCTTA 4755

QY 4636 -----GAAAAAGCAGAGCTGATGTCTAACTG 4662  
Db 4756 GAATCAGTTTGTGCGAGTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGTCTAACTG 4815  
QY 4663 CTTTGGAAATCCCTGCTGAAACTGGAAGGTGATGACCGCTAGACATGAATGCACGCTG 4722  
Db 4816 CTTTGGAAATCCCTGCTGAAACTGGAAGGTGATGACCGCTAGACATGAATGCACGCTG 4875  
QY 4723 CCTTTCAGTACAGAGTGTGTTTGGTGGGCAACGAGAAAGGGTCTACGCGCTGAATGTC 4782  
Db 4876 CCTTTCAGTACAGAGTGTGTTTGGTGGGCAACGAGAAAGGGTCTACGCGCTGAATGTC 4935  
QY 4783 TTGAAAAATCCCTTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAATTTATATATC 4842  
Db 4936 TTGAAAAATCCCTTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAATTTATATATC 4995  
QY 4843 AAGGACCTGAGAGGCTACTCATATAGCAGGAGAAAGAGCGGCACATGTGTCTTTGGAC 4902  
Db 4996 AAGGACCTGAGAGGCTACTCATATAGCAGGAGAAAGAGCGGCACATGTGTCTTTGGAC 5055  
QY 4903 GTGAGAAATGTAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGAGCATCTCA 4962  
Db 5056 GTGAGAAATGTAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGAGCATCTCA 5115  
QY 4963 CCACAAATTTTGAAGTGTCAAAGGGCTGCCACTTGTTTGGGCGAGCAAGATTGAGAAC 5022  
Db 5116 CCACAAATTTTGAAGTGTCAAAGGGCTGCCACTTGTTTGGGCGAGCAAGATTGAGAAC 5175  
QY 5023 GGGCTCTGATCTGTGAGCCATGCCAGCAAAAGTGTGCTATTTCTCGCTACAAAGAAAC 5082  
Db 5176 GGGCTCTGATCTGTGAGCCATGCCAGCAAAAGTGTGCTATTTCTCGCTACAAAGAAAC 5235  
QY 5083 CTGAGCAATATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5142  
Db 5236 CTGAGCAATATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5295  
QY 5143 TTCAACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTTACGAAATTCGATGAAGCAG 5202  
Db 5296 TTCAACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTTACGAAATTCGATGAAGCAG 5355  
QY 5203 TACAGCTCGAGGAATTCCTGGATGAAGATGACCAATTCCTTGGACCTGCTGTGTTGCC 5262  
Db 5356 TACAGCTCGAGGAATTCCTGGATGAAGATGACCAATTCCTTGGACCTGCTGTGTTGCC 5415  
QY 5363 GCCTTCTTCCAAAGCTTCCCTGTCTCAATCGTCAGGTGAACGCGCAGGCGAGCGAGAG 5322  
Db 5416 GCCTTCTTCCAAAGCTTCCCTGTCTCAATCGTCAGGTGAACGCGCAGGCGAGCGAGAG 5475  
QY 5323 GAGTACTTGTGTGTTTCCACGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5382  
Db 5476 GAGTACTTGTGTGTTTCCACGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5535  
QY 5383 CGCACAGACGATCTCAAGTGGAGTTCGCTTACCTTTGGCCCTTTGCCCTACAGAGAACCTAT 5442  
Db 5536 CGCACAGACGATCTCAAGTGGAGTTCGCTTACCTTTGGCCCTTTGCCCTACAGAGAACCTAT 5595  
QY 5443 CTGTGTTGTGACCACTTCAACTCACTCGAAGTAAATTTGAGATCCAGGCAACGCTCTCAGCA 5502  
Db 5596 CTGTGTTGTGACCACTTCAACTCACTCGAAGTAAATTTGAGATCCAGGCAACGCTCTCAGCA 5655  
QY 5503 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAACCCCGCGCTACCTGGGCCCTGCCATT 5562  
Db 5656 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAACCCCGCGCTACCTGGGCCCTGCCATT 5715  
QY 5563 TCCTCAGAGCGATTTACTTTGGGCTCCTCATACAGGATAAATTTAAGGGTCAATTTGCTGC 5622  
Db 5716 TCCTCAGAGCGATTTACTTTGGGCTCCTCATACAGGATAAATTTAAGGGTCAATTTGCTGC 5775  
QY 5623 AAGGAAAACTGTGTAAGGAGTCCGGACATGAAACACCCGGGCCCGCTCCACTCCCGC 5682  
Db 5776 AAGGAAAACTGTGTAAGGAGTCCGGACATGAAACACCCGGGCCCGCTCCACTCCCGC 5835  
QY 5683 AGCAGCCCCAACAGCAGGCGCCACCCACCGTACAAACGAGCACATCACCAAGCGGTGCC 5742

Db 5836 AGCAGCCCAACAGGAGGAGCCCAACCACTGTAACAGGACACATCAACAGCGGTGGCC 5895  
Qy 5743 TCAGCCAGCGCGCCGCGAGGAGCCAGCCAGCCGAGAGGCAAGCAACCCCAACCCG 5802  
Db 5896 TCCAGCCAGCGCGCCGCGAGGAGCCAGCCAGCCGAGAGGCAAGCAACCCCAACCCG 5955  
Qy 5803 TACCGAGGCGCGGACCGAGCTGCGCAGGAGCAAGTCTCTGCGCGCCCGCTGGAGCA 5862  
Db 5956 TACCGAGGCGCGGACCGAGCTGCGCAGGAGCAAGTCTCTGCGCGCCCGCTGGAGCA 6015  
Qy 5863 GAGAGTCCCGCGCGGATGTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 5922  
Db 6016 GAGAGTCCCGCGCGGATGTCAGCAGCGGAGAGCGGTCCCGCGGAGGCTGTTT 6075  
Qy 5923 GAGAGCAGCAGGCGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCAGGAG 5982  
Db 6076 GAGAGCAGCAGGCGCGGCTGCTGCGGAGCGGTGAGGACCCCGCTGTCAGGAG 6135  
Qy 5983 AACAGGAGAGGCGAGAGTGC 6005  
Db 6136 AACAGGCTGTGGACCACTTTC 6158

## RESULT 7

US-10-262-511-1  
; Sequence 1, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Mallet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Eilerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Blina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; PRIOR FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038

; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: Curasequid version 0.1  
; SEQ ID NO 1  
; LENGTH: 6189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(6159)  
US-10-262-511-1  
Query Match 91.4%; Score 5631; DB 13; Length 6189;  
Best Local Similarity 95.3%; Pred. No. 0;  
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;  
Qy 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGCTGTGAACCCATT 60  
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGCTGTGAACCCATT 60  
Qy 61 GCCAGCGGGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTATGACTCAA 120  
Db 61 GCCAGCGGGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTATGACTCAA 120  
Qy 121 CAGCAGATGTCTCTCTTCCGAGAGGAGGATATTAGATCCCTCTTGTCTCTTTGAA 180  
Db 121 CAGCAGATGTCTCTCTTCCGAGAGGAGGATATTAGATCCCTCTTGTCTCTTTGAA 180  
Qy 181 GAATGAGTGAAGTGTCTCTGATGAAGATTAAGACGCTGAGCACTTTGTCGGAAGTAT 240  
Db 181 GAATGAGTGAAGTGTCTCTGATGAAGATTAAGACGCTGAGCACTTTGTCGGAAGTAT 240  
Qy 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAGAGACTTCGGAAGTCAGA 300  
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAGAGACTTCGGAAGTCAGA 300  
Qy 301 AGTCTTCTAGGTTGTGTCACCTTTGCTGAAGTGCAGTGGTAAGAGAGAAAGCAACCGGG 360  
Db 301 AGTCTTCTAGGTTGTGTCACCTTTGCTGAAGTGCAGTGGTAAGAGAGAAAGCAACCGGG 360  
Qy 361 GACATCTATGCTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
Db 361 GACATCTATGCTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
Qy 421 TTTTTCAGGAGAGCGGAAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Db 421 TTTTTCAGGAGAGCGGAAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Qy 481 CAGTATGCTTTCAGGAG 540  
Db 481 CAGTATGCTTTCAGGAG 540  
Qy 541 GACTTGTCTGTCACCTTTTGAATAGATATGAGGACCAAGTGAAGAGAGAGAGAGAGAG 600  
Db 541 GACTTGTCTGTCACCTTTTGAATAGATATGAGGACCAAGTGAAGAGAGAGAGAGAGAG 600  
Qy 601 TACCTAGCTGAGTGAATTTGGCTGTTTACAGGCTTCACTGATGGATACGTCATCGA 660  
Db 601 TACCTAGCTGAGTGAATTTGGCTGTTTACAGGCTTCACTGATGGATACGTCATCGA 660  
Qy 661 GACATCAAGCTGAG 720  
Db 661 GACATCAAGCTGAG 720

Qy	721	GGATCTGCCGCGAAAAATGAATTCAACAAGATGGTGAATGCCAAATCTCCCGATTGGGACC	780
Db	721	GGATCTGCCGCGAAAAATGAATTCAACAACA-----GGTGAATGCCAAATCTCCCGATTGGGACC	777
Qy	781	CCAGATTACATGCTCTCTGAAGTCTGACTGTGATGAAACGGGATCGAAAGACCACTAC	840
Db	778	CCAGATTACATGCTCTCTGAAGTCTGACTGTGATGAAACGGGATCGAAAGACCACTAC	837
Qy	841	GGCTTGAATGTGACTGTGCTCAGTGGCGTGATTCCTATGAGATGATTTATGGGAGA	900
Db	838	GGCTTGAATGTGACTGTGCTCAGTGGCGTGATTCCTATGAGATGATTTATGGGAGA	897
Qy	901	TCCCCCTTCGACAGGGAACTCTGCCAGAACCTTCAATAACAATATGAAATTTCCAGCGG	960
Db	898	TCCCCCTTCGACAGGGAACTCTGCCAGAACCTTCAATAACAATATGAAATTTCCAGCGG	957
Qy	961	TTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTATCTGATTCATAAAGC	1020
Db	958	TTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTATCTGATTCATAAAGC	1017
Qy	1021	TTGTTGTGGCGGCAGAAAGAGACCTGAAGTTTGAAGTCTTTTCTGCTCATCTCTTCTTC	1080
Db	1018	TTGTTGTGGCGGCAGAAAGAGAGACCTGAAGTTTGAAGTCTTTTCTGCTCATCTCTTCTTC	1077
Qy	1081	TCTAAAATTGACTTGGAAACAATTTCGTAACTCTCTCTCCGCCCTTCGTTCCTCCACCTCAAG	1140
Db	1078	TCTAAAATTGACTTGGAAACAATTTCGTAACTCTCTCTCCGCCCTTCGTTCCTCCACCTCAAG	1137
Qy	1141	TCTGACGATGACACTCTCAAATTTTGTATGAACACGAGAGAATTCGTGGGTTTCATCTCT	1200
Db	1138	TCTGACGATGACACTCTCAAATTTTGTATGAACACGAGAGAATTCGTGGGTTTCATCTCT	1197
Qy	1201	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGGTTTTCG	1260
Db	1198	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGGTTTTCG	1257
Qy	1261	TACAGCAAGGCACTGGGGATCTTTGGTGAATCTGAGTCTGTTGTGCGGTCGTGACTCC	1320
Db	1258	TACAGCAAGGCACTGGGGATCTTTGGTGAATCTGAGTCTGTTGTGCGGTCGTGACTCC	1317
Qy	1321	CCTGCCAAGACTAGCTCCATGGAAAGAAACTTCTCATCAAAACGAAGAGCTACAAGAC	1380
Db	1318	CCTGCCAAGACTAGCTCCATGGAAAGAAACTTCTCATCAAAACGAAGAGCTACAAGAC	1377
Qy	1381	TCTCAGGACAAGTGTCACAAGATGGAGCAGGAATGACCCGGTTACATCGGAGAGTGCA	1440
Db	1378	TCTCAGGACAAGTGTCACAAGATGGAGCAGGAATGACCCGGTTACATCGGAGAGTGCA	1437
Qy	1441	GAGGTGAGGCTGTGCTTAGTCTCAGAGGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA	1500
Db	1438	GAGGTGAGGCTGTGCTTAGTCTCAGAGGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA	1497
Qy	1501	TCCCTCTCTGGACGAGGACCTTGCTACCTACATCACAGAAATGCAGTAGCTTAAAGCGAAAGT	1560
Db	1498	TCCCTCTCTGGACGAGGACCTTGCTACCTACATCACAGAAATGCAGTAGCTTAAAGCGAAAGT	1557
Qy	1561	TTGAGCAAGCACGGATGAGGTTGTCACGAGGAGTGATCAAAAGCACTGACGCTTCTCCAT	1620
Db	1558	TTGAGCAAGCACGGATGAGGTTGTCACGAGGAGTGATCAAAAGCACTGACGCTTCTCCAT	1617
Qy	1621	GATATCAGAGACGAGCGGAACTCCAAAGAAATCAAGACGAGGAGTACCAGGCTCAA	1680
Db	1618	GATATCAGAGACGAGCGGAACTCCAAAGAAATCAAGACGAGGAGTACCAGGCTCAA	1677
Qy	1681	GTGAAAGAAATGAGGTTGATGATGAATCAGTTGGAAAGGATCTTGTCTCAGCAAGAGA	1740
Db	1678	GTGAAAGAAATGAGGTTGATGATGAATCAGTTGGAAAGGATCTTGTCTCAGCAAGAGA	1737
Qy	1741	CGGAGTGATCTCTACGAATCTGAGCTCAGAGAGTCTCGGCTTGCTGCTGAAGAATTCAG	1800
Db	1738	CGGAGTGATCTCTACGAATCTGAGCTCAGAGAGTCTCGGCTTGCTGCTGAAGAATTCAG	1797
Qy	1801	CGGAAAGCGACAGAATCTCAGCATAAACTCTGTGAAGCTTAAGGATCAAGGGAAGCCTGAA	1860

1798	CGGAAAGCGCAGAAATGTCAGCATAAACTGTTGAAGCTTAAGATCAGGGGAAGCCTGAA	1857
1861	GTGGGAAATATCGGAAACTGGAAAGATCAATGCTGCAGCAGCAGCTCAAAATTCAGGAG	1920
1858	GTGGGAAATATCGGAAACTGGAAAGATCAATGCTGCAGCAGCAGCTCAAAATTCAGGAG	1917
1921	CTCCAGAGAAACTGGAGAGGCTG-----	1945
1918	CTCCAGAGAACTGGAGAGGCTGTAAAGCCAGCAGGAGGCCACCGAGCTGCTGCAG	1977
1946	-----CAAAGGACCGAGCCGAGCGGAGCTGGAGAAGCTGCAGAACCGAGAG	1992
1978	AATATCGCCAGCAAAGAGCGAGCCGAGAGGAGCTGAGAAAGCTGCAGAACCGAGAG	2037
1993	GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGGAAAGCCGCCAATTCCTG	2052
2038	GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAGCTGAGGAAAGCCGCCAATTCCTG	2097
2053	GAGAAACAGGTAAAGAGACTAGAGACCATGGAGCGGTAGAGAAACAGACTGAAAGGATGAC	2112
2098	GAGNAACAGGTAAAGAGACTAGAGACCATGGAGCGGTAGAGAAACAGACTGAAAGGATGAC	2157
2113	ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGTGAGCTCGAAGAG	2172
2158	ATCCAGACAAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCGTGAGCTCGAAGAG	2217
2173	AAACATCGGAGGCCAAAGCTCAGCCACGACCTAGAGCTGACCTGAAACAGAAAGAG	2232
2218	AAACATCGGAGGCCAAAGTCTCAGCCAGCAGCTAGAGTGCACCTGAAACAGAAAGAG	2277
2233	CAGCACTATCAGGAAAGATTTAAAGTGTTCGACAATCAGATAAAGAAAGACCTGGCTGAC	2292
2278	CAGCACTATCAGGAAAGATTTAAAGTATTTCGACAATCAGATAAAGAAAGACCTGGCTGAC	2337
2293	AAGAGACATCTGAGAACATGTCGAGACACGAGGAGGAGGCCCATGAGAGGGCCAAA	2352
2338	AAGAGACATCTGAGAACATGTCGAGACACGAGGAGGAGGCCCATGAGAGGGCCAAA	2397
2353	ATTTCTCAGGAAACAGAGGGCGATGATCAATGCTATGGATTTCAAAGATCAGATCCCTGGAA	2412
2398	ATTTCTCAGGAAACAGAGGGCGATGATCAATGCTATGGATTTCAAAGATCAGATCCCTGGAA	2457
2413	CAGAGGATTTGTGNACTGTCTGAAGCCCAATAACTTGCAGCAATAGCAGTCTTTTACC	2472
2458	CAGAGGATTTGTGNACTGTCTGAAGCCCAATAACTTGCAGCAATAGCAGTCTTTTACC	2517
2473	CAAAAGGAACATGAAGGCCCAAGAAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC	2532
2518	CAAAAGGAACATGAAGGCCCAAGAAGAGATGATTTCTGAACTCAGGCCAACAGAAATTTTAC	2577
2533	CTGAGACACAGGCTGGGAAAGTTGGAGGCCAGAAACCGAATACTGAGGAGCAGCTGGAG	2592
2578	CTGAGACACAGGCTGGGAAAGTTGGAGGCCAGAAACCGAATACTGAGGAGCAGCTGGAG	2637
2593	AAGATCAGCCACCAAGACCAACAGTGCAGAAAGTTCGGCTGCTGAACTGGAGACAAGATTG	2652
2638	AAGATCAGCCACCAAGACCAACAGTGCAGAAAGTTCGGCTGCTGAACTGGAGACAAGATTG	2697
2653	CGGAGGTCAGTCTTAGAGCACGAGGAGCGAAACTGGAGCTCAAGGCCAGCTCACAGAG	2712
2698	CGGAGGTCAGTCTTAGAGCACGAGGAGCGAAACTGGAGCTCAAGGCCAGCTCACAGAG	2757
2713	CTACAGCTCTCCCTGCGAGGAGCGGAGTCACAGTCAGCCCTGCAAGGCTGCACGGCGG	2772
2758	CTACAGCTCTCCCTGCGAGGAGCGGAGTCACAGTCAGCCCTGCAAGGCTGCACGGCGG	2817
2773	GCCCTGGAGAGCCAGCTTCGCCAGGCCAAGACAGAGCTGGAAGAGACCACAGCAGAGCT	2832
2818	GCCCTGGAGAGCCAGCTTCGCCAGGCCAAGACAGAGCTGGAAGAGACCACAGCAGAGCT	2877
2833	GAAGAGAGATCCAGGCACCTCAGGCCACATAGAGATGAATTCAGCGCCTAATTTGATGCT	2892



QY 4903 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGCGACATCTCA 4962  
Db 5053 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGCGACATCTCA 5112  
QY 4963 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACCTGTGTTGGGGCAGGCAAGATTGAGAAC 5022  
Db 5113 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACCTGTGTTGGGGCAGGCAAGATTGAGAAC 5172  
QY 5023 GGGCTCTGCATCTGTGAGCCATGCCCCAGCAAGTGTCTATCTCCGCTACAAAGAAAC 5082  
Db 5173 GGGCTCTGCATCTGTGAGCCATGCCCCAGCAAGTGTCTATCTCCGCTACAAAGAAAC 5232  
QY 5083 CTCACCAATATCTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5142  
Db 5233 CTCACCAATATCTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5292  
QY 5143 TTCACCAATATCTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5202  
Db 5293 TTCACCAATATCTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCCAC 5352  
QY 5203 TACAGCTCGAGGAATCTCGATAGAAATGACCAATCTTGGCACCCTGCTGTGTTGCC 5262  
Db 5353 TACAGCTCGAGGAATCTCGATAGAAATGACCAATCTTGGCACCCTGCTGTGTTGCC 5412  
QY 5263 GCCTCTTCCAAAGCTTCCCTGCTCAATCGTGCAGTGAACAGCGGCGAGCGAGAG 5322  
Db 5413 GCCTCTTCCAAAGCTTCCCTGCTCAATCGTGCAGTGAACAGCGGCGAGCGAGAG 5472  
QY 5323 GAGTACTTGTCTGTGTTCCAGCAATTTGGAGTGTTCGTGGATCTTAGGAGAGCTGAC 5382  
Db 5473 GAGTACTTGTCTGTGTTCCAGCAATTTGGAGTGTTCGTGGATCTTAGGAGAGCTGAC 5532  
QY 5383 CCACAGAGCACTCAAGTGGAGTCCCTTACCTTTGGGCTTTGCCCTACAGAGAACCTTAT 5442  
Db 5533 CCACAGAGCACTCAAGTGGAGTCCCTTACCTTTGGGCTTTGCCCTACAGAGAACCTTAT 5592  
QY 5443 CTGTTTGTGACCCACTTCAACTCACTGGAAGTAATGAGATCCAGGACGCTCTCAGCA 5502  
Db 5593 CTGTTTGTGACCCACTTCAACTCACTGGAAGTAATGAGATCCAGGACGCTCTCAGCA 5652  
QY 5503 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAACCCGCGTACTCGGCCCTGCCATT 5562  
Db 5653 GGGACCCCTGCGGAGCGTACCTGGACATCCCGAACCCGCGTACTCGGCCCTGCCATT 5712  
QY 5563 TCCTCAGGAGCGATTACTTGGCGTCTCATACAGGATAAAATTAAGGTCAATTTGCTGC 5622  
Db 5713 TCCTCAGGAGCGATTACTTGGCGTCTCATACAGGATAAAATTAAGGTCAATTTGCTGC 5772  
QY 5623 AAGGGAACCTCGTGAAGAGTCCGGCACTGAACCCAGCCGCGGCGCTCCACCTCCCGC 5682  
Db 5773 AAGGGAACCTCGTGAAGAGTCCGGCACTGAACCCAGCCGCGGCGCTCCACCTCCCGC 5832  
QY 5683 AGCAGCCCAACAAAGGAGGCGCCACCCAGTACACGAGCACATCAACGAGCGCGTGCC 5742  
Db 5833 AGCAGCCCAACAAAGGAGGCGCCACCCAGTACACGAGCACATCAACGAGCGCGTGCC 5892  
QY 5743 TCAGCCCAAGCGCCGCGAAAGCCCGACGACCCGCGAGAGCCAGCACCCACCGC 5802  
Db 5893 TCAGCCCAAGCGCCGCGAAAGCCCGACGACCCCGCGAGAGCCAGCACCCACCGC 5952  
QY 5803 TACCGGAGGGGCGGACCGAGTGGCAGGGACAAGTCTCTGGCGCGCCCTGGAGCGA 5862  
Db 5953 TACCGGAGGGGCGGACCGAGTGGCAGGGACAAGTCTCTGGCGCGCCCTGGAGCGA 6012  
QY 5863 GAGAAGTCCCGCGCGGAGTGTCTCAGACCGGAGAGAGCGGTTCCTCCCGGAGGCTGTTT 5922  
Db 6013 GAGAAGTCCCGCGCGGAGTGTCTCAGACCGGAGAGAGCGGTTCCTCCCGGAGGCTGTTT 6072  
QY 5923 GAAGACAGCAGAGGGGCGCGCTGCTCGGGAGCCGCTGAGGACCCGCTGTCAGAGTG 5982  
Db 6073 GAAGACAGCAGAGGGGCGCGCTGCTCGGGAGCCGCTGAGGACCCGCTGTCAGAGTG 6132  
QY 5983 AACAGGGAAGAGGGCAGAGTGCTCTCAAGTTTTCACGTTTAACTGTCACCT 6037

Db 6133 AACAGGTGTGGGACGAGTCTTTCAGTATAAATCTCAGCAGAAAAACCAACTCCT 6187  
RESULT 8  
US-09-964-956-10  
; Sequence 10, Application US/09964956  
; Publication No. US20040043926A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: Macdougall, John R  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David  
; APPLICANT: Gunther, Erik  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Leach, Martin D  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20040043926A1 Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-124  
; CURRENT APPLICATION NUMBER: US/09/964,956  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/235,631  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235,633  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/235,808  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,064  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,065  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,066  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/236,135  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/237,434  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/238,321  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,399  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,396  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/276,667  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/294,823  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: 60/304,868  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 6189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-964-956-10  
Query Match 91.4%; Score 5631; DB 13; Length 6189;  
Best Local Similarity 95.3%; Pred. No. 0;  
Matches 5944; Conservative 0; Mismatches 45; Indels 246; Gaps 4;  
QY 1 ATGTGTGAGTCAATATGGAGCGCGGAATCCCTTTGATGCTGTGCTGAACCCATT 60  
Db 1 ATGTGTGAGTCAATATGGAGCGCGGAATCCCTTTGATGCTGTGCTGAACCCATT 60

QY	61	GCACGCCGGCCTCAGGCTGAACTGTCTTCCTCAGGGGAAACCAACCTTTATGACTCAA	120
DB	61	GCACGCCGGCCTCAGGCTGAACTGTCTTCCTCAGGGGAAACCAACCTTTATGACTCAA	120
QY	121	CAGCAGATGTCTCCTCTTTCCGACAAGGGATATTAGATGCCCTCTTTGTCTCTTTGAA	180
DB	121	CAGCAGATGTCTCCTCTTTCCGAGAAGGGATATTAGATGCCCTCTTTGTCTCTTTGAA	180
QY	181	GAATCAGTCAGCCTGCTCTGATGAAGATTAAAGCAGTGAGCAACTTTGTCGGGAAGTAT	240
DB	181	GAATCAGTCAGCCTGCTCTGATGAAGATTAAAGCAGTGAGCAACTTTGTCGGGAAGTAT	240
QY	241	TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCNAAGGACTCTGGAAGTCAGA	300
DB	241	TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCNAAGGACTCTGGAAGTCAGA	300
QY	301	AGTCTGTGAGTGTGGTGCACCTTTGCTCAAGTCGAGTGGTAAAGAGAAAGCAACCGGG	360
DB	301	AGTCTGTGAGTGTGGTGCACCTTTGCTCAAGTCGAGTGGTAAAGAGAAAGCAACCGGG	360
QY	361	GACATCTATGCTATGAAGTGAATGAAGAAGAGGCTTTATTGGCCAGAGGAGGTTTCA	420
DB	361	GACATCTATGCTATGAAGTGAATGAAGAAGAGGCTTTATTGGCCAGAGGAGGTTTCA	420
QY	421	TTTTTTGAGGAAGCGGAAACATATTATCTCGAAGCAAGCCCGTGGATCCCCAATTA	480
DB	421	TTTTTTGAGGAAGCGGAAACATATTATCTCGAAGCAAGCCCGTGGATCCCCAATTA	480
QY	481	CAGTATGCTTTCCAGCAAAAATCACTTTATCTGATGGAGGAATATCAGCCTGGAGGG	540
DB	481	CAGTATGCTTTCCAGCAAAAATCACTTTATCTGATGGAGGAATATCAGCCTGGAGGG	540
QY	541	GACTTGCTGCACATTTTGAAATAGATATGAGGACAGTTAGATGAACCTTGATACAGTTT	600
DB	541	GACTTGCTGCACATTTTGAAATAGATATGAGGACAGTTAGATGAACCTTGATACAGTTT	600
QY	601	TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCACTGATGGGATAGTCGCATCGA	660
DB	601	TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCACTGATGGGATAGTCGCATCGG	660
QY	661	GACATCAAGCCTGAGACATCTCGTTGACCGCAGAGACACATCAAGCTGGTGGATTTT	720
DB	661	GACATCAAGCCTGAGACATCTCGTTGACCGCAGAGACACATCAAGCTGGTGGATTTT	720
QY	721	GGATCTGCCGGAAAAATGAAATCAAAACAAGATGCTGAATGCCAACTCCCGATTGGGACC	780
DB	721	GGATCTGCCGGAAAAATGAAATCAAAACA--GGTGAATGCCAACTCCCGATTGGGACC	777
QY	781	CCAGATTACATGGCTCTGTAAGTCTGACTGTGTGATGAACGGGGATGGAAAGGCACCTAC	840
DB	778	CCAGATTACATGGCTCTGTAAGTCTGACTGTGTGATGAACGGGGATGGAAAGGCACCTAC	837
QY	841	GGCCTGACCTGTGACTGGTGTGAGTGGGCGTGATTCCTATGAGATGATTTATGGGAGA	900
DB	838	GGCCTGACCTGTGACTGGTGTGAGTGGGCGTGATTCCTATGAGATGATTTATGGGAGA	897
QY	901	TCGCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATTAACATTTGAAATTTCCAGCGG	960
DB	898	TCGCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATTAACATTTGAAATTTCCAGCGG	957
QY	961	TTTTTTGAAATTTCCAGATGACCCCAAGTGAGCAGTGCACCTTTCTTGATCTGATTCAAAGC	1020
DB	958	TTTTTTGAAATTTCCAGATGACCCCAAGTGAGCAGTGCACCTTTCTTGATCTGATTCAAAGC	1017
QY	1021	TTGTTGTGCGGCCAGAAAGAGACTGAAGTTTGAAGGCTTTTGCTGCCATCCTTTCTTC	1080
DB	1018	TTGTTGTGCGGCCAGAAAGAGACTGAAGTTTGAAGGCTTTTGCTGCCATCCTTTCTTC	1077
QY	1081	TCTAAAATTCAGCTGGACACATTCGTAACCTCTCCGCCCTTCGTTCCACGCTCAAG	1140
DB	1078	TCTAAAATTCAGCTGGAAACAATTCGTAAACGCTCTCCGCCCTTCGTTCCACGCTCAAG	1137
QY	1141	TCTGACGATGACACCTCCAAATTTTGATGAACACAGAGAAGAAATTCGTGGGTTTTCATCTCT	1200

1138	Db	TCTGACGATGACAGCTCCAAATTTTGTATGAACACAGAGAAGAAATTCGTGGGTTCATCCTCT	1197
1201	Qy	CCGTGCAGCTGAGCCCTCAGGCTTCTCGSGTGAAGAACTGCCTGTGTGGGGTTTTCG	1260
1198	Db	CCGTGCAGCTGAGCCCTCAGGCTTCTCGSGTGAAGAACTGCCTGTGTGGGGTTTTCG	1257
1261	Qy	TACAGCAAGCACTGGGGATTCTTGGTAGATCTCAGTCTGTGTGTGGGTCTGGACTCC	1320
1258	Db	TACAGCAAGCACTGGGGATTCTTGGTAGATCTCAGTCTGTGTGTGGGTCTGGACTCC	1317
1321	Qy	CTTCCCAAGACTAGCTCCATGGAAAGAACTTCTCATCAAGCAAGAGCTTACAGAC	1380
1318	Db	CTTCCCAAGACTAGCTCCATGGAAAGAACTTCTCATCAAGCAAGAGCTTACAGAC	1377
1381	Qy	TCTCAGCACAAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTACATCGGAGGTGTCA	1440
1378	Db	TCTCAGCACAAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTACATCGGAGGTGTCA	1437
1441	Qy	GAGGTGAGAGCTGTCTTAGTCAAGAGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA	1500
1438	Db	GAGGTGAGAGCTGTCTTAGTCAAGAGAGGTGGAGCTGAAGCCCTCTGAGACTCAGAGA	1497
1501	Qy	TCCCTCTCGAGCAGGACCTTGTCTACTCATCACAGAATCGATAGCTTAAAGCGAAGT	1560
1498	Db	TCCCTCTCGAGCAGGACCTTGTCTACTCATCACAGAATCGATAGCTTAAAGCGAAGT	1557
1561	Qy	TTGAGCAAGCA CGGATGGAGTGTCTCCAGGAGGATGACAAAGCACTGCAGTTCCTCAT	1620
1558	Db	TTGAGCAAGCA CGGATGGAGTGTCTCCAGGAGGATGACAAAGCACTGCAGTTCCTCAT	1617
1621	Qy	GATATCAGACAGCAGAGCCGGAAGCTTCCAAGAAATCAAGACGAGAGTACCAGGCTCAA	1680
1618	Db	GATATCAGACAGCAGAGCCGGAAGCTTCCAAGAAATCAAGACGAGAGTACCAGGCTCAA	1677
1681	Qy	GTGGAAGAAATGAGGTGTGATGAAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA	1740
1678	Db	GTGGAAGAAATGAGGTGTGATGAAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA	1737
1741	Qy	CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAATTCAG	1800
1738	Db	CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAATTCAG	1797
1801	Qy	CGGAAAGCGACAGAAATCTCAGCATAAACTCTTGAAGCTTAAGGATCAAGGGAAGCTCAA	1860
1798	Db	CGGAAAGCGACAGAAATCTCAGCATAAACTCTTGAAGCTTAAGGATCAAGGGAAGCTCAA	1857
1861	Qy	GTGGGGAATATCGGAAAATGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG	1920
1858	Db	GTGGGGAATATCGGAAAATGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG	1917
1921	Qy	CTCAAGAGAACTGGAGAGGCTG-----	1945
1918	Db	CTCAAGAGAACTGGAGAGGCTGTTAAAGCCGACCGAGGCCACCGAGCTGTGCAG	1977
1946	Qy	-----CNAAGAGCGAGCCGAGAGGGAGCTGGAGAGCTGCAGAACTGCAGAACCGAGAG	1992
1978	Db	AATATCCGCGAGCAAGAGCGAGCCGAGAGGGAGCTGGAGAACTGCAGAACCGAGAG	2037
1993	Qy	GATCTTCTGAAGGCATCAGAAGAGCTGTGGAGCTTGGAGACGCGCCATTCCTGTG	2052
2038	Db	GATCTTCTGAAGGCATCAGAAGAGCTGTGGAGCTTGGAGACGCGCCATTCCTGTG	2097
2053	Qy	GAGAACAAAGTAAAGAGACTAGAGACCAATGGAGCGGTAGAGAAAACAGACTGAAGGATGAC	2112
2098	Db	GAGAACAAAGTAAAGAGACTAGAGACCAATGGAGCGGTAGAGAAAACAGACTGAAGGATGAC	2157
2113	Qy	ATCCAGACAAAATCCCAAGATCCAGAGATGCGTGTATTAATTTCTGGAGCTCGAAGAG	2172
2158	Db	ATCCAGACAAAATCCCAAGATCCAGAGATGCGTGTATTAATTTCTGGAGCTCGAAGAG	2217
2173	Qy	AAACATCGGAGGCCCAAGCTCTCAGCCGACCACTAGAGGTGCACCTGAAACAGAAAGAG	2232

Db 2218 AAACATCGGAGGCCCCAAGTCTCAGCCGAGCACCTAGAAAGTGCACCTGAAAACAGAAAGAG 2277  
Qy 2233 CAGCACTATGAGGAAAAGATTAAAGTGTGGACAATCAGATAAAAGAAAGACTGCTGTCAC 2292  
Db 2278 CAGCACTATGAGGAAAGATTAAAGTATTGGACAATCAGATAAAAGAAAGACTGCTGTCAC 2337  
Qy 2293 AAGGAGACACTGGAGAAACATGATGAGAGACACGAGGAGGAGGCCCATGAGAAAGGGCAAA 2352  
Db 2338 AAGGAGACACTGGAGAAACATGATGAGAGACACGAGGAGGAGGCCCATGAGAAAGGGCAAA 2397  
Qy 2353 ATTCTCAGCGAACAGAAAGGCGATGATCAATGCTATGGATTTCAAGATCAGATCCCTGGAA 2412  
Db 2398 ATTCTCAGCGAACAGAAAGGCGATGATCAATGCTATGGATTTCAAGATCAGATCCCTGGAA 2457  
Qy 2413 CAGAGGATTTGGAACACTGCTGAAAGCAATAAACCTTGAGGCAAAATPAGCAGTCTTTTACC 2472  
Db 2458 CAGAGGATTTGGAACACTGCTGAAAGCAATAAACCTTGAGGCAAAATPAGCAGTCTTTTACC 2517  
Qy 2473 CAAAGGACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532  
Db 2518 CAAAGGACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577  
Qy 2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCGGAAACTCGAGGAGCAGCTGGAG 2592  
Db 2578 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCGGAAACTCGAGGAGCAGCTGGAG 2637  
Qy 2593 AAGATCAGCCACCAAGACACACAGTACAGAAATCGCTGCGAACTGGAGACTGGAGACAGATTG 2652  
Db 2638 AAGATCAGCCACCAAGACACACAGTACAGAAATCGCTGCGAACTGGAGACTGGAGACAGATTG 2697  
Qy 2653 CGGGAGGTCAGTCTAGAGCAGGAGCAGAGAACTGGAGCTCAAGCGCCAGCTCACAGAG 2712  
Db 2698 CGGGAGGTCAGTCTAGAGCAGGAGCAGAGAACTGGAGCTCAAGCGCCAGCTCACAGAG 2757  
Qy 2713 CTACAGCTCTCCCTGAGAGGCGCAGTCAAGTTGACAGCCCTGAGGCTGCACGGGCG 2772  
Db 2758 CTACAGCTCTCCCTGAGAGGCGCAGTCAAGTTGACAGCCCTGAGGCTGCACGGGCG 2817  
Qy 2773 GGCCTGGAGAGCCAGTCTCCAGCGGAGAGACAGAGCTGGAAGAGACACACAGCAAGCT 2832  
Db 2818 GGCCTGGAGAGCCAGTCTCCAGCGGAGAGACAGAGCTGGAAGAGACACACAGCAAGCT 2877  
Qy 2833 GAAGAGGAGATCCAGGCACTCAGGCACATAGAGATGAAATCAGCGCAAAATTTGATGCT 2892  
Db 2878 GAAGAGGAGATCCAGGCACTCAGGCACATAGAGATGAAATCAGCGCAAAATTTGATGCT 2937  
Qy 2893 CTTCTGTAACAGCTGTACTGTAATCAGACCTGGAGGACAGCTAAACCGAGTGCAGG 2952  
Db 2938 CTTCTGTAACAGCTGTACTGTAATCAGACCTGGAGGACAGCTAAACCGAGTGCAGG 2997  
Qy 2953 GACACGCTGAACTCAACCAACCAAACTTCTACTTTGTCCAAACAACTCGATGAGGCTTCT 3012  
Db 2998 GACACGCTGAACTCAACCAACCAAACTTCTACTTTGTCCAAACAACTCGATGAGGCTTCT 3057  
Qy 3013 GCGGCCAACGACAGATTGTAACTGCGAAGTGAAGTGAACATCTCCGCGGGAGATC 3072  
Db 3058 GCGGCCAACGACAGATTGTAACTGCGAAGTGAAGTGAACATCTCCGCGGGAGATC 3117  
Qy 3073 ACCGAAAGAGATGAGCTTACCAGCAGAGCAACCAAGATGAGGCTCTGAAAGACCAAG 3132  
Db 3118 ACCGAAAGAGATGAGCTTACCAGCAGAGCAACCAAGATGAGGCTCTGAAAGACCAAG 3177  
Qy 3133 TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTAAACGATCAGCTGTAGAA 3192  
Db 3178 TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTAAACGATCAGCTGTAGAA 3237  
Qy 3193 AAAGAGGGCAGTGGAGGCTGGAGGCGTCTGGGTGATGAGAAATCCAGATTGAG 3252  
Db 3238 AAAGAGGGCAGTGGAGGCGTCTGGGTGATGAGAAATCCAGATTGAG 3297  
Qy 3253 TGTCCGGTTCCAGAGCTGCAGAGAAATGCTGGACACCGAGAAAACAGAGCAGGCGAGAGCC 3312  
Db 3298 TGTCCGGTTCCAGAGCTGCAGAGAAATGCTGGACACCGAGAAAACAGAGCAGGCGAGAGCC 3357

Qy 3313 GATCAGCGGATCACCAGTCTCCGACGGTGGAGCTGCGAGTGAAGAGCAGCAGGCT 3372  
Db 3358 GATCAGCGGATCACCAGTCTCCGACGGTGGAGCTGCGAGTGAAGAGCAGCAGGCT 3417  
Qy 3373 GAGATTCTCCCTCTGACGAGGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCCTCTCT 3432  
Db 3418 GAGATTCTCCCTCTGACGAGGCTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCCTCTCT 3477  
Qy 3433 GACAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAATGCCGAAGCTTA 3492  
Db 3478 GACAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAATGCCGAAGCTTA 3537  
Qy 3493 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGCAGCA 3552  
Db 3538 CAGCAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGCAGCA 3597  
Qy 3553 TTACAGCAGCAGATGGACCTGCAAGAAAATCAATTTTCGCTCTGACTCAAGGACTGCAA 3612  
Db 3598 TTACAGCAGCAGATGGACCTGCAAGAAAATCAATTTTCGCTCTGACTCAAGGACTGCAA 3657  
Qy 3613 GAAGCTCTAGATCGGCGTGTACTCTGAAGCAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGA 3672  
Db 3658 GAAGCTCTAGATCGGCGTGTACTCTGAAGCAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGA 3717  
Qy 3673 GAAAACATTCAGGTTCTCTATTCTCATGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3732  
Db 3718 GAAAACATTCAGGTTCTCTATTCTCATGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3777  
Qy 3733 CAAACCAAACTCATGATTTTCTGCAAGCAAAATGGAACCACTGCTTAAAGAAAAG 3792  
Db 3778 CAAACCAAACTCATGATTTTCTGCAAGCAAAATGGAACCACTGCTTAAAGAAAAG -- 3835  
Qy 3793 GGTTTATTAGTCGACGAGAAAGAGGACCTGCTTTTACCCACACAGGTTCTCTCAGTAC 3852  
Db 3836 -----AGGTGCTCTCAGTAC 3852  
Qy 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTAGAGAGGCC 3912  
Db 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTAGAGAGGCC 3912  
Qy 3913 CTTTCAAGAGACCCGATCGAGCTCCGCTCCGCGGGAGGAGCTGCCACCGCAAGCA 3972  
Db 3913 CTTTCAAGAGACCCGATCGAGCTCCGCTCCGCGGGAGGAGCTGCCACCGCAAGCA 3972  
Qy 3973 ACGGACCAACCAACCCATCCACCGCAGCCACCGCAGGAGCAGAGATCGCCATCTCCGC 4032  
Db 3973 ACGGACCAACCAACCCATCCACCGCAGCCACCGCAGGAGCAGAGATCGCCATCTCCGC 4032  
Qy 4033 ATCGTGGGTGCGCAGAGCAGCCAGCTGCGATGAGCTGCTGGCGCCGCGCATCCAGC 4092  
Db 4033 ATCGTGGGTGCGCAGAGCAGCCAGCTGCGATGAGCTGCTGGCGCCGCGCATCCAGC 4092  
Qy 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGGTCTTAAAGAACCGATGCAC 4152  
Db 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGAAATTTAGTCGGGTCTTAAAGAACCGATGCAC 4152  
Qy 4153 CACAATATTCCTCACCAGTCAACGTAAGACTGAACATGCGAGGCCACAAAGTGTGCTGTG 4212  
Db 4153 CACAATATTCCTCACCAGTCAACGTAAGACTGAACATGCGAGGCCACAAAGTGTGCTGTG 4212  
Qy 4213 TGTCTGGATACCGTGCATTTGGAGCGCAGGCAATCCAAATGTCTCAGATGTTCAGGTGATG 4272  
Db 4213 TGTCTGGATACCGTGCATTTGGAGCGCAGGCAATCCAAATGTCTCAGATGTTCAGGTGATG 4272  
Qy 4273 TGTCAACCCCAAGTGTCCACGCTGCTTGCAGGCCACCTCGCGCTTGCCTGTGTAATATGCC 4332  
Db 4273 TGTCAACCCCAAGTGTCCACGCTGCTTGCAGGCCACCTCGCGCTTGCCTGTGTAATATGCC 4332  
Qy 4333 ACACATTCACCGAGGCGCTTCTGCGGTGAACAAATGAACCTCCCGAGGCTTCCAGACCAAG 4392  
Db 4333 ACACATTCACCGAGGCGCTTCTGCGGTGAACAAATGAACCTCCCGAGGCTTCCAGACCAAG 4392

QY 4393 GAGCCAGCAGCAGCTTGCACCTGGAAGGCTGATGAGGTGCCAGGATATCAACAACGA 4452  
Db 4393 GAGCCAGCAGCAGCTTGCACCTGGAAGGCTGATGAGGTGCCAGGATATCAACAACGA 4452  
QY 4453 GGACAGCAAGCTGGGACAGAGATCAATTTCTCGAGAGGATCAAAAGTCCTATTAT 4512  
Db 4453 GGACAGCAAGCTGGGACAGAGATCAATTTCTCGAGAGGATCAAAAGTCCTATTAT 4512  
QY 4513 GACAAATCAAGCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCCCTCC 4572  
Db 4513 GACAAATCAAGCAGAGAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCCCTCC 4572  
QY 4573 GAGGGGATGTATCTATTTCATGTGCGTGGTGGTTCGGAATTCGCAAAATACAGCAAA 4632  
Db 4573 GAGGGGATGTATCTATTTCATGTGCGTGGTGGTTCGGAATTCGCAAAATACAGCAAA 4632  
QY 4633 GCA----- 4635  
Db 4633 GCAGATGCCATACATATCTGAAGATGGAATCTCACCCGACACACCCTGCTGGCCCGGG 4692  
QY 4636 ----- 4635  
Db 4693 AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4752  
QY 4636 -----GAAAAAGCAGAGCTGATGCTTAACTG 4662  
Db 4753 GAATCAGTTGTCGACGCTGGAGAGTTTCTAGGGAAGAAAGCAGAGCTGATGCTTAACTG 4812  
QY 4663 CTTGGAAATCCCTGCTGAAATCGGAAGGTGATGACCGTCTAGACATGAATCGACGCTG 4722  
Db 4813 CTTGGAAATCCCTGCTGAAATCGGAAGGTGATGACCGTCTAGACATGAATCGACGCTG 4872  
QY 4723 CCTTCAGTGACACAGGTGTTGTTGGGACCGAGGAAGGGCTCTACCCCTGAATGTC 4782  
Db 4873 CCTTCAGTGACACAGGTGTTGTTGGGACCGAGGAAGGGCTCTACCCCTGAATGTC 4932  
QY 4783 TTGAAAAATCCCTAACCATGTCCAGGAATGGAGCAGTCTTCCAAATTTATATTATC 4842  
Db 4933 TTGAAAAATCCCTAACCATGTCCAGGAATGGAGCAGTCTTCCAAATTTATATTATC 4992  
QY 4843 AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCGACTGTGCTTGTGGAC 4902  
Db 4993 AAGGACCTGGAGAGCTACTCATGATAGCAGGTGAAGAGCGGCGACTGTGCTTGTGGAC 5052  
QY 4903 GTCAAGAAAGTGAACAGTCCCTGGCCGACGTCACCTGCTGCCAGCCGACATCTCA 4962  
Db 5053 GTCAAGAAAGTGAACAGTCCCTGGCCGACGTCACCTGCTGCCAGCCGACATCTCA 5112  
QY 4963 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5022  
Db 5113 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGCAAGATTGAGAAC 5172  
QY 5023 GGGCTCTGATCTGTCAGCCATGCCAGCAAAAGTGTCAATTTCTCGCTACAAAGAAAC 5082  
Db 5173 GGGCTCTGATCTGTCAGCCATGCCAGCAAAAGTGTCAATTTCTCGCTACAAAGAAAC 5232  
QY 5083 CTCAGCAATATCTGCATCCGAAAGAGATAGAGACTCAGAGCCCTGAGCTGTATCCAC 5142  
Db 5233 CTCAGCAATATCTGCATCCGAAAGAGATAGAGACTCAGAGCCCTGAGCTGTATCCAC 5292  
QY 5143 TTCACAAATACAGTATCTCTATGGAACCAATAATTCTACGAATCGACATGAAGCAG 5202  
Db 5293 TTCACAAATACAGTATCTCTATGGAACCAATAATTCTACGAATCGACATGAAGCAG 5352  
QY 5203 TACAGCTCGAGGAATTCCTGATAGATGACCATCTCTGGACCTGCTGTGTTGCC 5262  
Db 5353 TACAGCTCGAGGAATTCCTGATAGATGACCATCTCTGGACCTGCTGTGTTGCC 5412  
QY 5263 GCTCTTCCACAGCTTCCCTGTCTCAATGTGAGGTGAACAGCGCAGGCGAGCAGAG 5322  
Db 5413 GCTCTTCCACAGCTTCCCTGTCTCAATGTGAGGTGAACAGCGCAGGCGAGCAGAG 5472  
QY 5323 GAGTACTTGTCTGTTTCCACGAATTTGGAGTGTTCGTGGATTCTTACGGAAGAGCTAGC 5382

Db 5473 GAGTACTTGTCTGTTTCCAGGAATTTGGAGTGTTCGTGGATTCTTACGGAAGAGCTAGC 5532  
QY 5383 GCACAGACGATCTCAAGTGGAGTGCCTTACTTTGGCTTTTGGCTACAGAGAACCCTAT 5442  
Db 5533 GCACAGACGATCTCAAGTGGAGTGCCTTACTTTGGCTTTTGGCTACAGAGAACCCTAT 5592  
QY 5443 CTGTTTGGAGCCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCTCAGCA 5502  
Db 5593 CTGTTTGGAGCCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCTCAGCA 5652  
QY 5503 GGGACCCCTGCGGAGCGTACTTGGACATCCGCAACCCGCGCTACCTTGGGCCCTTGCATT 5562  
Db 5653 GGGACCCCTGCGGAGCGTACTTGGACATCCGCAACCCGCGCTACCTTGGGCCCTTGCATT 5712  
QY 5563 TCCTCAGAGGATTTACTTTGGCTGCTCATACAGGATTAATTAAGGGTCAATTTGCTGC 5622  
Db 5713 TCCTCAGAGGATTTACTTTGGCTGCTCATACAGGATTAATTAAGGGTCAATTTGCTGC 5772  
QY 5623 AAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACCCAGGCGGCCCTCCACCTCCCGC 5682  
Db 5773 AAGGAAACCTCGTGAAGGAGTCCGGCACTGAACACCCAGGCGGCCCTCCACCTCCCGC 5832  
QY 5683 AGCAGCCCAACAGAGAGGCGCCACCGTACAAAGAGACATCAACAGCGGCTGGCC 5742  
Db 5833 AGCAGCCCAACAGAGAGGCGCCACCGTACAAAGAGACATCAACAGCGGCTGGCC 5892  
QY 5743 TCCAGCCAGCGCGCCGAGGCGCCAGCAGCCAGGAGAGCAAGCTCTCTGGCCGCCCTTGGAGCGA 5802  
Db 5893 TCCAGCCAGCGCGCCGAGGCGCCAGGAGAGCAAGCTCTCTGGCCGCCCTTGGAGCGA 5952  
QY 5803 TACCGAGGAGGCGGACCGAGCTGCGCAGGGAAGAAGTCTCTGGCCGCCCTTGGAGCGA 5862  
Db 5953 TACCGAGGAGGCGGACCGAGCTGCGCAGGGAAGAAGTCTCTGGCCGCCCTTGGAGCGA 6012  
QY 5863 GAGAAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCCTCCCGGAGGCTGTTT 5922  
Db 6013 GAGAAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCCTCCCGGAGGCTGTTT 6072  
QY 5923 GAAGACAGCAGCGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGAGTG 5982  
Db 6073 GAAGACAGCAGCGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGTCCAGAGTG 6132  
QY 5983 AACAGGAGAGGCGAGAGTGCCTCTCAGTGTTCACGTTTAACTGTCACT 6037  
Db 6133 AACAGGAGAGGCGAGAGTGCCTCTCAGTGTTCACGTTTAACTGTCACT 6187

## RESULT 9

US-10-262-511-13  
; Sequence 13, Application US/10262511  
; Publication No. US2004003823A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Elberman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.

APPLICANT: Rastelli, Luca  
APPLICANT: Stone, David J.  
APPLICANT: Pena, Carol E. A.  
APPLICANT: Shenoy, Suresh G.  
APPLICANT: Shimkets, Richard A.  
APPLICANT: Rothenberg, Mark E.  
APPLICANT: Leach, Martin D.  
APPLICANT: Agee, Michele L.  
APPLICANT: Berghs, Constance  
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: 21402-462C  
CURRENT APPLICATION NUMBER: US/10/262,511  
CURRENT FILING DATE: 2003-05-28  
PRIORITY APPLICATION NUMBER: 60/326,483  
PRIORITY FILING DATE: 2001-10-02  
PRIORITY APPLICATION NUMBER: 60/373,815  
PRIORITY FILING DATE: 2002-04-19  
PRIORITY APPLICATION NUMBER: 60/327,917  
PRIORITY FILING DATE: 2001-10-09  
PRIORITY APPLICATION NUMBER: 60/381,642  
PRIORITY FILING DATE: 2002-05-17  
PRIORITY APPLICATION NUMBER: 60/328,029  
PRIORITY FILING DATE: 2002-10-09  
PRIORITY APPLICATION NUMBER: 60/381,038  
PRIORITY FILING DATE: 2002-05-16  
PRIORITY APPLICATION NUMBER: 60/328,056  
PRIORITY FILING DATE: 2001-10-09  
PRIORITY APPLICATION NUMBER: 60/373,260  
PRIORITY FILING DATE: 2002-04-17  
PRIORITY APPLICATION NUMBER: 60/373,826  
PRIORITY FILING DATE: 2002-04-19  
PRIORITY APPLICATION NUMBER: 60/327,435  
PRIORITY FILING DATE: 2001-10-05  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 439  
SOFTWARE: CuroSeqList version 0.1  
SEQ ID NO 13  
LENGTH: 6201  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(6198)  
US-10-262-511-13

Query March  
Best Local Similarity 91.4%; Score 5629.8; DB 13; Length 6201;  
Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;  
1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAAACCCATT 60  
1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATGCTGGTCTGCTGAAACCCATT 60  
61 GCCAGCGGGCTCCAGGCTGATCTGTTCCAGGGGAACCCACCTTTATGACTCAA 120  
61 GCCAGCGGGCTCCAGGCTGATCTGTTCCAGGGGAACCCACCTTTATGACTCAA 120  
121 CAGCAGATGCTCCTCTTTCCCGAGAGGATATAGATGCCCTTTCTCTTTGAA 180  
121 CAGCAGATGCTCCTCTTTCCCGAGAGGATATAGATGCCCTTTCTCTTTGAA 180  
181 GAATGAGTCCAGCTGCTCTGATGAAGATTAAGCAGTGAGCACTTTCGCGAGATAT 240  
181 GAATGAGTCCAGCTGCTCTGATGAAGATTAAGCAGTGAGCACTTTCGCGAGATAT 240  
241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGGACTTCGAAGTCAG 300  
241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCCGCAAGGACTTCGAAGTCAG 300  
301 AGTCTTGAGCTTGCTGCTACTTTGCTGAAGTGCAGGTGATAGAGAGAAAGCAACCGGG 360  
301 AGTCTTGAGCTTGCTGCTACTTTGCTGAAGTGCAGGTGATAGAGAGAAAGCAACCGGG 360

QY 361 GACATCTATGCTATGAAAGTGTATGAAGAAAGAGGCTTTTATGGCCAGAGAGAGGTTTCA 420  
Db |||||  
QY 361 GACATCTATGCTATGAAAGTGTATGAAGAAAGAGGCTTTTATGGCCAGAGAGGTTTCA 420  
Db |||||  
QY 421 TTTTGTGAGAAAGAGCGGAACATATTTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Db |||||  
QY 421 TTTTGTGAGAAAGAGCGGAACATATTTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Db |||||  
QY 481 CAGTATGCCCTTTCCAGGACAAAATACCTTTATCTGATGAGGAATATCAGCTCGAGGG 540  
Db |||||  
QY 481 CAGTATGCCCTTTCCAGGACAAAATACCTTTATCTGATGAGGAATATCAGCTCGAGGG 540  
Db |||||  
QY 541 GACTTCTGCTCACTTTTGAATAGATAGAGGACCACTTAGATGAGAAAACCTGTACAGTTT 600  
Db |||||  
QY 541 GACTTCTGCTCACTTTTGAATAGATAGAGGACCACTTAGATGAGAAAACCTGTACAGTTT 600  
Db |||||  
QY 601 TACTAGCTGAGCTGATTTTGGCTGTTCCAGCGTTTCACTGATGGGATACGTGCAATCGA 660  
Db |||||  
QY 601 TACTAGCTGAGCTGATTTTGGCTGTTCCAGCGTTTCACTGATGGGATACGTGCAATCGG 660  
Db |||||  
QY 661 GACATCAAGCCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTTGATTTT 720  
Db |||||  
QY 661 GACATCAAGCCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTTGATTTT 720  
Db |||||  
QY 721 GGATCTCCCGGAAATGAATTCAAAACAGATGCTGAATGCCAACTCCCGATTGGGACC 780  
Db |||||  
QY 721 GGATCTCCCGGAAATGAATTCAAAACAGATGCTGAATGCCAACTCCCGATTGGGACC 777  
Db |||||  
QY 781 CCAGATTACATGCTCCTGAAGTGTGATGAAACCGGGATGAAAGGACACCTAC 840  
Db |||||  
QY 778 CCAGATTACATGCTCCTGAAGTGTGATGAAACCGGGATGAAAGGACACCTAC 837  
Db |||||  
QY 841 GGCTGGAAGTGTGATGCTGCTGAGTGGGCTGATTCCTTATGAGATGATTTTGGGAGA 900  
Db |||||  
QY 838 GGCTGGAAGTGTGATGCTGCTGAGTGGGCTGATTCCTTATGAGATGATTTTGGGAGA 897  
Db |||||  
QY 901 TCCCTCTCCAGAGGGAACCTCTGCCAGAACCTTCAATAACATTTATGAATTTCCAGCGG 960  
Db |||||  
QY 898 TCCCTCTCCAGAGGGAACCTCTGCCAGAACCTTCAATAACATTTATGAATTTCCAGCGG 957  
Db |||||  
QY 961 TTTTGTGAATTTCCAGATGACCCCAAGTGTGAGCAGTGAATTTCTGATGATGATCAAGC 1020  
Db |||||  
QY 958 TTTTGTGAATTTCCAGATGACCCCAAGTGTGAGCAGTGAATTTCTGATGATGATCAAGC 1017  
Db |||||  
QY 1021 TTGTTGTGGCCAGAAAGAGAGACTGAAATTTGAAGTCTTTGCTGCCATCTTTCTTC 1080  
Db |||||  
QY 1018 TTGTTGTGGCCAGAAAGAGAGACTGAAATTTGAAGTCTTTGCTGCCATCTTTCTTC 1077  
Db |||||  
QY 1081 TCTAAAATTTGACTGGAACAAACATTTGTAATCTCTCTCCCTCCCTTCCTCCACCTCAAG 1140  
Db |||||  
QY 1078 TCTAAAATTTGACTGGAACAAACATTTGTAATCTCTCTCCCTTCCTCCACCTCAAG 1137  
Db |||||  
QY 1141 TCTCAAGATGACCTCCAAATTTTGTATGAACACAGAGAGAAATTCGTGGGTTTCACTCT 1200  
Db |||||  
QY 1138 TCTCAAGATGACCTCCAAATTTTGTATGAACACAGAGAGAAATTCGTGGGTTTCACTCT 1197  
Db |||||  
QY 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTGTGGGGTTTTTCG 1260  
Db |||||  
QY 1198 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTGTGGGGTTTTTCG 1257  
Db |||||  
QY 1261 TACAGCAAGGCACTGGGATTTTGTGATGATCTGTTGTGTTGTTGTTGTTGTTGTTGTT 1320  
Db |||||  
QY 1258 TACAGCAAGGCACTGGGATTTTGTGATGATCTGTTGTGTTGTTGTTGTTGTTGTTGTT 1317  
Db |||||  
QY 1321 CTGCAAGATGACTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380  
Db |||||  
QY 1318 CTGCAAGATGACTCCATGGAAGAAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1377  
Db |||||  
QY 1381 TCTCAGGACAAAGTGTCAAGATGAGCAGGAAATGACCCGGTTTACATCGGAGAGTGTCA 1440  
Db |||||  
QY 1378 TCTCAGGACAAAGTGTCAAGATGAGCAGGAAATGACCCGGTTTACATCGGAGAGTGTCA 1437  
Db |||||  
QY 1441 GAGGTGAGGCTGTGCTTAGTTCAGAAAGGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500

Db 1438 GAGGTGAGGCTGTGCTTTAGTCAGAGGAGGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1497  
QY 1501 TCCTCTCTGGAGCAGGACCTTGTCTACCTACATCACAGAATGCGAGTGTAAAGCGAAGT 1560  
Db 1498 TCCTCTCTGGAGCAGGACCTTGTCTACCTACATCACAGAATGCGAGTGTAAAGCGAAGT 1557  
QY 1561 TTGGAGCAAGCAGGATGAGGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1620  
Db 1558 TTGGAGCAAGCAGGATGAGGTGTCCAGGAGGATGACAAAGCACTGAGCTTCTCCAT 1617  
QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGGCTCAA 1680  
Db 1618 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGGCTCAA 1677  
QY 1681 GTGGAAAGAAATGAGTTGATGAAATCAGTTGGAAAGAGATCTTGTCTCAGCAAGAAGA 1740  
Db 1678 GTGGAAAGAAATGAGTTGATGAAATCAGTTGGAAAGAGATCTTGTCTCAGCAAGAAGA 1737  
QY 1741 CGGAGTGATCTTACGAATCTGAGCTGAGAGAGTCTCGCTTCTGCTGARGAATTCAG 1800  
Db 1738 CGGAGTGATCTTACGAATCTGAGCTGAGAGAGTCTCGCTTCTGCTGARGAATTCAG 1797  
QY 1801 CGGAAAGCGACAGAAATGTGAGCAATAAATCTTTGAAGGCTTAAGGATCAAGGAGGCTCAA 1860  
Db 1798 CGGAAAGCGACAGAAATGTGAGCAATAAATCTTTGAAGGCTTAAGGATCAGGGAAGCTGAA 1857  
QY 1861 GTGGGAGAAATATCGGAAATCTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
Db 1858 GTGGGAGAAATATCGGAAATCTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1917  
QY 1921 CTCGAAGAGAAATCTGGAGAGGCT----- 1944  
Db 1918 CTCGAAGAGAAATCTGGAGAGGCTGTAAAGCCAGCAGCGGAGCCACCGAGCTGTGCGAG 1977  
QY 1945 -----GCAAGAGCGAGCCGAGAGGAGCTGGAGAACTGCGAGAACCGAGAG 1992  
Db 1978 AATATCCGCGCAGCAAGAGAGCGAGCCGAGAGGAGCTGGAGAACTGCGAGAACCGAGAG 2037  
QY 1993 GATTTCTTCTGAAGGCAATCAGAAAGAGCTGTGGAGCTGAGGAACTGAGGAACTGCTG 2052  
Db 2038 GATTTCTTCTGAAGGCAATCAGAAAGAGCTGTGGAGCTGAGGAACTGAGGAACTGCTG 2097  
QY 2053 GAGAACAGGTAAGAGACTAGAGACCTAGGAGCCTAGAGAAACAGACTGAAGGATGAC 2112  
Db 2098 GAGAACAGGTAAGAGACTAGAGACCTAGGAGCCTAGAGAAACAGACTGAAGGATGAC 2157  
QY 2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGCGCTGATTAATTTCTGGAGCTCGAAGAG 2172  
Db 2158 ATCCAGACAAATCCCAACAGATCCAGCAGATGCGCTGATTAATTTCTGGAGCTCGAAGAG 2217  
QY 2173 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCCTAGAAAGTGCACTGAAACAGAAAGAG 2232  
Db 2218 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCCTAGAAAGTGCACTGAAACAGAAAGAG 2277  
QY 2233 CAGCACTATGAGGAAAGATTAAGTGTGGCAGTGTGAGCAGTGAAGAGAGAGAGAGAGAG 2292  
Db 2278 CAGCACTATGAGGAAAGATTAAGTGTGGCAGTGTGAGCAGTGAAGAGAGAGAGAGAGAG 2337  
QY 2293 AAGGAGACACTGAGAAATCATGATGAGAGACACAGAGGAGGAGGCCATGAGAGAGGCAAA 2352  
Db 2338 AAGGAGACACTGAGAAATCATGATGAGAGACACAGAGGAGGAGGCCATGAGAGAGGCAAA 2397  
QY 2353 ATTCTCAGGAAACAGAGCGGATGATCAATGCTATGAGTTCCAGATCAGATCCTCGAA 2412  
Db 2398 ATTCTCAGGAAACAGAGCGGATGATCAATGCTATGAGTTCCAGATCAGATCCTCGAA 2457  
QY 2413 CAGAGGATTTGGAACTGTCTGAAGCCCAATAAATCTTGAGCAAAATAGCAGTCTTTTACC 2472  
Db 2458 CAGAGGATTTGGAACTGTCTGAAGCCCAATAAATCTTGAGCAAAATAGCAGTCTTTTACC 2517  
QY 2473 CAAAGGAACTGAGGCCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2532

Db 2518 CAAAGGAACTGAGGCCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2577  
QY 2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG 2592  
Db 2578 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG 2637  
QY 2593 AAGATCAGCCACCAAGACCAAGTGAAGAATCGGCTGTGGAACTGGAGACAAAGATTG 2652  
Db 2638 AAGATCAGCCACCAAGACCAAGTGAAGAATCGGCTGTGGAACTGGAGACAAAGATTG 2697  
QY 2653 CGGAGGCTCAGCTTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCCGAGCTCAGAGAG 2712  
Db 2698 CGGAGGCTCAGCTTAGAGCAGAGGAGCAGAACTGGAGCTCAAGCCGAGCTCAGAGAG 2757  
QY 2713 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCTTGACAGCTTGACAGCGGCG 2772  
Db 2758 CTACAGCTCTCCCTGCGAGGAGCGAGTCAAGTTGACAGCTTGACAGCTTGACAGCGGCG 2817  
QY 2773 GCCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTTGAAGAGACACACAGCAAGACT 2832  
Db 2818 GCCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTTGAAGAGACACACAGCAAGACT 2877  
QY 2833 GAAGAGGAGATCCAGGCACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2892  
Db 2878 GAAGAGGAGATCCAGGCACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2937  
QY 2893 CTTGCTTAACAGCTGTACTGTAAATCAGAGCTTGAGAGGAGCAGCTTAAACAGCTGACCGAG 2952  
Db 2938 CTTGCTTAACAGCTGTACTGTGATCAGAGCTTGAGAGGAGCAGCTTAAACAGCTGACCGAG 2997  
QY 2953 GACAACTGCTGAACTCAACAAACAAACTTCTACTTGTCCAAACAACTCGATGAGGTTCT 3012  
Db 2998 GACAACTGCTGAACTCAACAAACAAACTTCTACTTGTCCAAACAACTCGATGAGGTTCT 3057  
QY 3013 GCGGCCAAACAGCAGATTTGACAACTGCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3072  
Db 3058 GCGGCCAAACAGCAGATTTGACAACTGCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3117  
QY 3073 ACCGAAACGAGAGATGAGCTTACCAGCAGAGCAACAGATGAGGCTCTGAGAGCAGAG 3132  
Db 3118 ACCGAAACGAGAGATGAGCTTACCAGCAGAGCAACAGATGAGGCTCTGAGAGCAGAG 3177  
QY 3133 TGCACTGCTGGAGAAACAGGCTCATGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3192  
Db 3178 TGCACTGCTGGAGAAACAGGCTCATGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3237  
QY 3193 AAAGAGCGGAGTGGAGGCGCTGGAGGAGGCTCTGGGAGCGCTGAGAAATCCCAAGTTGAG 3252  
Db 3238 AAAGAGCGGAGTGGAGGCGCTGGAGGAGGCTCTGGGAGCGCTGAGAAATCCCAAGTTGAG 3297  
QY 3253 TGTGCGGTTGAGAGCTGAGAGATGCTTGGACACCGAGAAACAGAGCAGGCGAGAGCC 3312  
Db 3298 TGTGCGGTTGAGAGCTGAGAGATGCTTGGACACCGAGAAACAGAGCAGGCGAGAGCC 3357  
QY 3313 GATCAGCGGATCAACGAGTCTGCGAGTGTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 3372  
Db 3358 GATCAGCGGATCAACGAGTCTGCGAGTGTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 3417  
QY 3373 GAGATTTCTGCTCTGAGCAGGCTCTCAAGAGCAGAGCTCAAGCCGAGAGCTCTCT 3432  
Db 3418 GAGATTTCTGCTCTGAGCAGGCTCTCAAGAGCAGAGCTCAAGCCGAGAGCTCTCT 3477  
QY 3433 GACAGCTCAATGACTGAGAGAGAGCATGTATGCTTTGAAATGAATGCCCGAGAGCTTA 3492  
Db 3478 GACAGCTCAATGACTGAGAGAGAGCATGTATGCTTTGAAATGAATGCCCGAGAGCTTA 3537  
QY 3493 CAGCAGAGCTGGAGACTGAACGAGGCTCAACAGAGCTTCTGAGAGCAAGCCCAA 3552  
Db 3538 CAGCAGAGCTGGAGACTGAACGAGGCTCAACAGAGCTTCTGAGAGCAAGCCCAA 3597  
QY 3553 TTACAGCAGCAGATGAGCTGAGAGAAATCAATCAATTTCCGCTGAGCTCAAGGACTGCAA 3612  
Db 3598 TTACAGCAGCAGATGAGCTGAGAGAAATCAATCAATTTCCGCTGAGCTCAAGGACTGCAA 3657

QY	3613	GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATGCTGGAGTATCAGCTG	3672	QY	4636	-----	4635
Db	3658	GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATGCTGGAGTATCAGCTG	3717	Db	4693	AGAACTCTACTGCTAGCTCCAGCTTCCCTGTGACAAACAGCGTGGGTCAACCGCTTAA	4752
QY	3673	GAAGCAATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGCACAATTTCTCAA	3732	QY	4636	-----	4662
Db	3718	GAAGCAATTCAGGTTCTCTATTCTCATGAAAGGTGAAATGGAAGCACAATTTCTCAA	3777	Db	4753	GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAAAGAGAGCTGATCTAACTG	4812
QY	3733	CAAAACCAATTCATGATTTTCTGCAAGCCAAATGACCACTGCTAAAGAAAAG	3792	QY	4663	CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGCTCTAGACATGAATCTGACGCTG	4722
Db	3778	CAAAACCAATTCATGATTTTCTGCAAGCCAAATGACCACTGCTAAAGAAAAG	3835	Db	4813	CTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGCTCTAGACATGAATCTGACGCTG	4872
QY	3793	GGTTTATTAGTCGACGGAAGAGACCCCTCTTTAACCACACAGAGTTCTCTCAGTAC	3852	QY	4723	CCCTTCAGTGACCAAGTGGTGTGGTGGACCGAGGAGGCTCTACGCCCTGAAATGTC	4782
Db	3836	-----	3852	Db	4873	CCCTTCAGTGACCAAGTGGTGTGGTGGACCGAGGAGGCTCTACGCCCTGAAATGTC	4932
QY	3853	AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTGCTGTGAGAGCTAGAGGAAGCC	3912	QY	4783	TTGAAAACTCCCTAAACCATGTCCAGCAATTTGGAGCAGTCTTCCAAATTTATATTATC	4842
Db	3853	AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTGCTGTGAGAGCTAGAGGAAGCC	3912	Db	4933	TTGAAAACTCCCTAAACCATGTCCAGCAATTTGGAGCAGTCTTCCAAATTTATATTATC	4992
QY	3913	CTTCAGAGACCCGATCGAGCTCGGTCCGCCGGGAGGAGCTGCCACCGCAAGCA	3972	QY	4843	AAGGACTCTGAGAGCTACTCATGATAGCAGAGAAAGAGCGGGCACTGTCTTTGGAC	4902
Db	3913	CTTCAGAGACCCGATCGAGCTCGGTCCGCCGGGAGGAGCTGCCACCGCAAGCA	3972	Db	4993	AAGGACTCTGAGAGCTACTCATGATAGCAGAGAAAGAGCGGGCACTGTCTTTGGAC	5052
QY	3973	ACGGACCAACCAACCCATCCACGCCAGCCACCGGAGGAGAGATGCCATGTCCGCC	4032	QY	4903	GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA	4962
Db	3973	ACGGACCAACCAACCCATCCACGCCAGCCACCGGAGGAGAGATGCCATGTCCGCC	4032	Db	5053	GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCGACATCTCA	5112
QY	4033	ATCGTCCGTCGACAGACACAGCCAGTCCGATGAGCTGCTGGCCCGCCATCCAGC	4092	QY	4963	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCAGGCAAGATTGAAAC	5022
Db	4033	ATCGTCCGTCGACAGACACAGCCAGTCCGATGAGCTGCTGGCCCGCCATCCAGC	4092	Db	5113	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCAGGCAAGATTGAAAC	5172
QY	4093	CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTCTTAAGGAACGATGCC	4152	QY	5023	GGCTCTGCATCTGTGACGCGCATGCCAGCAAGTCTGATCTTCCGCTACAAACGAAAC	5082
Db	4093	CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTCTTAAGGAACGATGCC	4152	Db	5173	GGCTCTGCATCTGTGACGCGCATGCCAGCAAGTCTGATCTTCCGCTACAAACGAAAC	5232
QY	4153	CACAAATTTCTCAACGATTCACGATGAGATGAGATGAGATGAGATGAGATGAGATGAG	4212	QY	5083	CTCAGCAAAATACCTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGACGTGTATCCAC	5142
Db	4153	CACAAATTTCTCAACGATTCACGATGAGATGAGATGAGATGAGATGAGATGAGATGAG	4212	Db	5233	CTCAGCAAAATACCTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGACGTGTATCCAC	5292
QY	4213	TGCTGGATACCGTGCACCTTGGACGCGAGGATCCAAATGTCTCGAATGTCAAGTGTATG	4272	QY	5143	TTCAACCAATTTACAGTATCTCAATTTGAAATCCCAATTTTACGAAATCGACATGAAGCAG	5202
Db	4213	TGCTGGATACCGTGCACCTTGGACGCGAGGATCCAAATGTCTCGAATGTCAAGTGTATG	4272	Db	5293	TTCAACCAATTTACAGTATCTCAATTTGAAATCCCAATTTTACGAAATCGACATGAAGCAG	5352
QY	4273	TGTCACCCCAAGTGTCCACGCTGCTTGCAGCCACCTCGGCTTCCCTGCTGAATATGCC	4332	QY	5203	TACACGCTCCAGGAAATTCCTGGATAGAAATGACCAATTCCTTGGCAGCTGTGTGTGGCC	5262
Db	4273	TGTCACCCCAAGTGTCCACGCTGCTTGCAGCCACCTCGGCTTCCCTGCTGAATATGCC	4332	Db	5353	TACACGCTCCAGGAAATTCCTGGATAGAAATGACCAATTCCTTGGCAGCTGTGTGTGGCC	5412
QY	4333	ACACACTTCACGAGGCTTCTGCGGTGACAAATGAACTCCCAAGGCTCCAGACCAAG	4392	QY	5263	GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACGCGAGGCGAGGAG	5322
Db	4333	ACACACTTCACGAGGCTTCTGCGGTGACAAATGAACTCCCAAGGCTCCAGACCAAG	4392	Db	5413	GCCTCTTCCAAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACGCGAGGCGAGGAG	5472
QY	4393	GAGCCAGCAGAGCTTGCACTGGAAGGTTGATGAAGGTGCCAGGAAATGAACAAACGA	4452	QY	5323	GAGTACTTGTGTGTTCACGAAATTTGGAGTGTTCGTGGATTTCTACGGAAGACGTAGC	5382
Db	4393	GAGCCAGCAGAGCTTGCACTGGAAGGTTGATGAAGGTGCCAGGAAATGAACAAACGA	4452	Db	5473	GAGTACTTGTGTGTTCACGAAATTTGGAGTGTTCGTGGATTTCTACGGAAGACGTAGC	5532
QY	4453	GGACAGCAAGGTTGGACAGGAGTACATGTTCTGGAGGATCAAAAGTCTCTATTAT	4512	QY	5383	CGCACAGACGATCTCAAGTGGAGTTCGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT	5442
Db	4453	GGACAGCAAGGTTGGACAGGAGTACATGTTCTGGAGGATCAAAAGTCTCTATTAT	4512	Db	5533	CGCACAGACGATCTCAAGTGGAGTTCGCTTACCTTTGGCCCTTTGCTTACAGAGAACCTTAT	5592
QY	4513	GACAAATGAAGCCAGAGAGCTGACAGAGCGGTTGGAGAAATTTGAGCTGTGCCCTTCC	4572	QY	5443	CTGTTTGTGACCACTTCAACTCACTCGAAGTAAATTGAGATCCAGGCGCTCTCTCAGCA	5502
Db	4513	GACAAATGAAGCCAGAGAGCTGACAGAGCGGTTGGAGAAATTTGAGCTGTGCCCTTCC	4572	Db	5593	CTGTTTGTGACCACTTCAACTCACTCGAAGTAAATTGAGATCCAGGCGCTCTCTCAGCA	5652
QY	4573	GACGGGATGTATCTATTATCATGTCGCTGCTGCTTCCGAACTGCGAAATACAGCCAA	4632	QY	5503	GGGACCCCTGCGGAGGCTTACCTGGACATCCCGAAACCCGGGCTACCTGGGCGCTTGCATT	5562
Db	4573	GACGGGATGTATCTATTATCATGTCGCTGCTGCTTCCGAACTGCGAAATACAGCCAA	4632	Db	5653	GGGACCCCTGCGGAGGCTTACCTGGACATCCCGAAACCCGGGCTACCTGGGCGCTTGCATT	5712
QY	4633	GCA-----	4635	QY	5563	TCCTCAGAGCGATTTACTTTGGCGCTCCTCATACAGGATAAATTAAAGGTCATTTGTGTC	5622
Db	4633	GCAGATGTCCCATACATACGAAAGTGAATCTACACCCGACACACCTGCTGGCCCGGG	4692	Db	5713	TCCTCAGAGCGATTTACTTTGGCGCTCCTCATACAGGATAAATTAAAGGTCATTTGTGTC	5772



778	CCAGATTACATGCTCCTGAAGTGCTGACTGTGTGATGAACGGGATGGAAAGGACCTAC	837
841	GGCCTGGACTGTGACTGTGGTGCAGTGGGGTGATGGCTATGAGATGATTTATGGGAGA	900
838	GGCCTGGACTGTGACTGTGGTGCAGTGGGGTGATGGCTATGAGATGATTTATGGGAGA	897
901	TTCCCCCTTCGCAGAGGGAACCTCTGCCAGAACCTTCAATAACAATATGAATTTCCAGCGG	960
898	TTCCCCCTTCGCAGAGGGAACCTCTGCCAGAACCTTCAATAACAATATGAATTTCCAGCGG	957
961	TTTTTTGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTTGATCTGATTCGAAGC	1020
958	TTTTTTGAAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTTGATCTGATTCGAAGC	1017
1021	TTCTGTGTGCGCCAGAAAAGAGACTGAAGTTTGAAGGTCCTTTGCTGCCATCTTTCTTC	1080
1018	TTGTGTGTGCGCCAGAAAAGAGACTGAAGTTTGAAGGTCCTTTGCTGCCATCTTTCTTC	1077
1081	TCATAAATTGACTGGAAACAATTCGTAAGTCTCTCTCCGCCCTTCGTTCCACCCCTCAAG	1140
1078	TCATAAATTGACTGGAAACAATTCGTAAGTCTCTCTCCGCCCTTCGTTCCACCCCTCAAG	1137
1141	TCGTGACGATCACACCTCCAAATTTTGATGAACACAGAGAAGAAATTCGTGGGTTTCATCCTCT	1200
1138	TCGTGACGATCACACCTCCAAATTTTGATGAACACAGAGAAGAAATTCGTGGGTTTCATCCTCT	1197
1201	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCCGTTTGTGGGGTTTTCG	1260
1198	CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCCGTTTGTGGGGTTTTCG	1257
1261	TACAGCAAGCACTGGGGATCTTGTGTAGATCTGAGTCTCTGTGTGCGGTCTGGACTCC	1320
1258	TACAGCAAGCACTGGGGATCTTGTGTAGATCTGAGTCTCTGTGTGCGGTCTGGACTCC	1317
1321	CCTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAAAGAAAGCTTACAAGAC	1380
1318	CCTGCCAAGACTAGCTCCATGGAAAAGAACTTCTCATCAAAAAGAAAGCTTACAAGAC	1377
1381	TCTCAGACAAAGTCTACAAGATGGACAGGAATGACCGGTTTACATCGGAGAGTGCA	1440
1378	TCTCAGACAAAGTCTACAAGATGGACAGGAATGACCGGTTTACATCGGAGAGTGCA	1437
1441	GAGGTGAGGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGGCCTCTGAGACTCAGAGA	1500
1438	GAGGTGAGGCTGTGCTTAGTCAAGAGGAGGTGGAGCTGAAGGCCTCTGAGACTCAGAGA	1497
1501	TCCCTCTCTGGACAGGACCTTGCTACCTACATCACAGAAATGCAGTACGTTTAAAGCCGAGT	1560
1498	TCCCTCTCTGGACAGGACCTTGCTACCTACATCACAGAAATGCAGTACGTTTAAAGCCGAGT	1557
1561	TTGAGCAAGCACGGATGGAGGTCTCCAGAGGATGACAAAGCACTGACAGCTTCTCCAT	1620
1558	TTGAGCAAGCACGGATGGAGGTCTCCAGAGGATGACAAAGCACTGACAGCTTCTCCAT	1617
1621	GATATCAGAGAGCAGACCGGAAGCTCCAGAAATCAAGACAGGAGTACCAAGGCTCAA	1680
1618	GATATCAGAGAGCAGACCGGAAGCTCCAGAAATCAAGACAGGAGTACCAAGGCTCAA	1677
1661	GTGGAAGAAATGAGGTTGATCATCAATCAGTTGGAAGAGGATCTGTCTCAGCAAGAAGA	1740
1678	GTGGAAGAAATGAGGTTGATCATCAATCAGTTGGAAGAGGATCTGTCTCAGCAAGAAGA	1737
1741	CGGAGTGAATCTTACGAATCTGAGCTGAGAGGTCCTGGCTTGTCTGTGAAGAAATTCAG	1800
1738	CGGAGTGAATCTTACGAATCTGAGCTGAGAGGTCCTGGCTTGTCTGTGAAGAAATTCAG	1797
1801	CGGAAAGCGACAGAAATGTCAGCAATAAATCTTTGAAGGCTAAGGATCAAGGAGCCTGAA	1860
1798	CGGAAAGCGACAGAAATGTCAGCAATAAATCTTTGAAGGCTAAGGATCAAGGAGCCTGAA	1857
1861	CTGGAGAAATATGCCAAACTCGAGAAGATCAATGCTGTGACAGCAGCTCAAAATTCAGAG	1920

1858	Db	GTGGGAATATATGCGAAACTTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG	1917
1921	Qy	CTCCACAGAGAAACTTGGAGAAAGGCT-----	1944
1918	Db	CTCCACAGAGAACTTGGAGAAAGGCTGTAAAAGCCAGCAGCGAGGCCACCGAGCTGCTCGAG	1977
1945	Qy	-----GCAAGGAGCGAGCGAGCGAGGAGCTGGAGAGCTGCGAGAACCTGCAGAACCGGAGAG	1992
1978	Db	AATATCGCGCAGGCCAAAGGACGAGCGAGGAGGCTGGAGAACTTGCAGAACCTGCAGAACCGGAGAG	2037
1993	Qy	GATTCCTTCTGAAGGCATCAGAAAAGAGCTGGTGGAGCTGAGGAAACCCGCCATTCTCTG	2052
2038	Db	GATTCCTTCTGAAGGCATCAGAAAAGAGCTGGTGGAGCTGAGGAAACCCGCCATTCTCTG	2097
2053	Qy	GAGAACAGGTAAAGAGACTAGAGACCATGGAGCGGTAGAGAAACAGACCTGAGAGGTAGC	2112
2098	Db	GAGAACAGGTAAAGAGACTAGAGACCATGGAGCGGTAGAGAAACAGACCTGAGAGGTAGC	2157
2113	Qy	ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCCTGGAGCTCGAAGAG	2172
2158	Db	ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTCCTGGAGCTCGAAGAG	2217
2173	Qy	AAAATCTGGGAGGCCAAGTCTGAGCCAGCACTAGAACTGCACCTGGAACAGAAAGAG	2232
2218	Db	AAAATCTGGGAGGCCAAGTCTGAGCCAGCACTAGAACTGCACCTGGAACAGAAAGAG	2277
2233	Qy	CAGCACTATGAGGAAAAAGATTTAAAGTGTGGACAAATCAGATAAAGAAAGACCTGGCTGAC	2292
2278	Db	CAGCACTATGAGGAAAAAGATTTAAAGTATTTGACAAATCAGATAAAGAAAGACCTGGCTGAC	2337
2293	Qy	AAGGAGACACTTGGAGAACATGATCCAGAGACACAGGAGGAGGCCCATGGAAGGCGCAA	2352
2338	Db	AAGGAGACACTTGGAGAACATGATCCAGAGACACAGGAGGAGGCCCATGGAAGGCGCAA	2397
2353	Qy	ATTCTCAGCAACAGAGAGCGCATGATCAATGCTATGGATTCCAAAGATCAGATCCCTGGAA	2412
2398	Db	ATTCTCAGCAACAGAGAGCGCATGATCAATGCTATGGATTCCAAAGATCAGATCCCTGGAA	2457
2413	Qy	CAGAGGATTTTGGAACTGTCTGAAGCCCAATAAATTCGAGCAATAGCAGTCTTTTACC	2472
2458	Db	CAGAGGATTTTGGAACTGTCTGAAGCCCAATAAATTCGAGCAATAGCAGTCTTTTACC	2517
2473	Qy	CAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC	2532
2518	Db	CAAGGAACATGAAGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC	2577
2533	Qy	CTGGAGACACAGGCTGGGAACTTGGAGGCCCAAGAACCGAAATCTGGAGGACAGCTGAG	2592
2578	Db	CTGGAGACACAGGCTGGGAACTTGGAGGCCCAAGAACCGAAATCTGGAGGACAGCTGAG	2637
2593	Qy	AAGATCAGCCACCAAGACCAACAGTGCACAGAACTCGGTCTGGAACCTGGAGACAGATTG	2652
2638	Db	AAGATCAGCCACCAAGACCAACAGTGCACAGAACTCGGTCTGGAACCTGGAGACAGATTG	2697
2653	Qy	CGGAGGTCAGTCTTAGAGCACGAGGAGCAGAAAACTGGAGCTCAAGCGCCAGCTCACAGAG	2712
2698	Db	CGGAGGTCAGTCTTAGAGCACGAGGAGCAGAAAACTGGAGCTCAAGCGCCAGCTCACAGAG	2757
2713	Qy	CTACAGCTCTCCCTGACGAGAGCGGAGTCAAGTTTGCAGCCCTCGAGGCTGCAAGGCG	2772
2758	Db	CTACAGCTCTCCCTGACGAGAGCGGAGTCAAGTTTGCAGCCCTCGAGGCTGCAAGGCG	2817
2773	Qy	GCCTCTGGAGAGCCAGCTTTCGCCAGCGCAAGACAGAGCTTGGAAAGAGACCACAGCAAGCT	2832
2818	Db	GCCCTGGAGAGCCAGCTTTCGCCAGCGCAAGACAGAGCTTGGAAAGAGACCACAGCAAGCT	2877
2833	Qy	GAAGAGGAGATCCAGGCACTACCGGCAATAGAGATGAAATTCAGCGCAAAATTTGATGCT	2892
2878	Db	GAAGAGGAGATCCAGGCACTACCGGCAATAGAGATGAAATTCAGCGCAAAATTTGATGCT	2937
2893	Qy	CTTCGTACAGCTGTACTGTGTAATCACAGACCTGGAGGAGCAGCTATAACAGCTGACCGAG	2952
2938	Db	CTTCGTACAGCTGTACTGTGTAATCACAGACCTGGAGGAGCAGCTATAACAGCTGACCGAG	2997





Qy	721	GGATCTGCCGGGAAAAATGAATTCAAACCAAGATGTGTAATCCCAACTCCCGATTGGGACC	780
Db	721	GGATCTGCCGGGAAAAATGAATTCAAACCAAGATGTGTAATCCCAACTCCCGATTGGGACC	780
Qy	781	CCAGATTACATGGCTCCCTGAAGTCTGCTACTCTGATGAACGGGGATGAAAAGGCACCTAC	840
Db	781	CCAGATTACATGGCTCCCTGAAGTCTGCTACTCTGATGAACGGGGATGAAAAGGCACCTAC	840
Qy	841	GGCCTGGACTGTGACTTGGTGTCTAGTGGGCGGTGATTGCCTATGAGATGCAATTTATGGGAGA	900
Db	841	GGCCTGGACTGTGACTTGGTGTCTAGTGGGCGGTGATTGCCTATGAGATGCAATTTATGGGAGA	900
Qy	901	TCCCCCTTCGCAGAGGGAACCTCTCCAGAACCTTCAATACATATGAATTTCCACGG	960
Db	901	TCCCCCTTCGCAGAGGGAACCTCTCCAGAACCTTCAATACATATGAATTTCCACGG	960
Qy	961	TTTTTGAATTTCCAGATGACCCCAAAGTGACAGTCACTTTCTTGATCTGTGATTCAAAGC	1020
Db	961	TTTTTGAATTTCCAGATGACCCCAAAGTGACAGTCACTTTCTTGATCTGTGATTCAAAGC	1020
Qy	1021	TTGTTGTGGCCCGAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATTCCTTCTTC	1080
Db	1021	TTGTTGTGGCCCGAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATTCCTTCTTC	1080
Qy	1081	TCATAAATTGACTGGAAACAACATTCTGTAACCTCCTCCCCCTTCGTTCCACCCCTCAAG	1140
Db	1081	TCATAAATTGACTGGAAACAACATTCTGTAACCTCCTCCCCCTTCGTTCCACCCCTCAAG	1140
Qy	1141	TCAGCAGTACACCTCCAAATTTTGTATGAACACAGAGAAGAATTCGTGGGTTTCATCCTCT	1200
Db	1141	TCAGCAGTACACCTCCAAATTTTGTATGAACACAGAGAAGAATTCGTGGGTTTCATCCTCT	1200
Qy	1201	CCGTGCCAGCTGAGCCCTCAGGCTTCTCCGGTCAAGAACTGCGCTTGTGGGGTTTTCG	1260
Db	1201	CCGTGCCAGCTGAGCCCTCAGGCTTCTCCGGTCAAGAACTGCGCTTGTGGGGTTTTCG	1260
Qy	1261	TACAGCAAGCCTCGGGATTCCTGGTAGATCTGAGTCTCTTGTGTCCGGTCTGGACTCC	1320
Db	1261	TACAGCAAGCCTCGGGATTCCTGGTAGATCTGAGTCTCTTGTGTCCGGTCTGGACTCC	1320
Qy	1321	CCGTGCCAGTACTCCATCGAAGAAACAACTTCTCATCAAAAGCAAGAGCTACAGAC	1380
Db	1321	CCGTGCCAGTACTCCATCGAAGAAACAACTTCTCATCAAAAGCAAGAGCTACAGAC	1380
Qy	1381	TCCTCAGCAAGAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTACATCGGAGTGTCA	1440
Db	1381	TCCTCAGCAAGAAGTGTCAAAAGATGGAGCAAGAAATGACCCGGTTACATCGGAGTGTCA	1440
Qy	1441	GAGGTGGAGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAGCCTCTGAGACTCAAGAGA	1500
Db	1441	GAGGTGGAGCTGTGCTTAGTTCAGAGGAGGTGGAGCTGAAGCCTCTGAGACTCAAGAGA	1500
Qy	1501	TCCCTCTCGAGCAGGACCTTGCTACTACATCACAGAATGCASTAGCTTAAAGCGAAGT	1560
Db	1501	TCCCTCTCGAGCAGGACCTTGCTACTACATCACAGAATGCASTAGCTTAAAGCGAAGT	1560
Qy	1561	TTGGAGCAAGCAGGATGGAGGTGTCCAGAGGATGACAAAGCACTGACAGTCTTCCAT	1620
Db	1561	TTGGAGCAAGCAGGATGGAGGTGTCCAGAGGATGACAAAGCACTGACAGTCTTCCAT	1620
Qy	1621	GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGAGTACCAGGCTCAA	1680
Db	1621	GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGAGTACCAGGCTCAA	1680
Qy	1681	GTGGAAGAAATGAGTTGATGATCAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA	1740
Db	1681	GTGGAAGAAATGAGTTGATGATCAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA	1740
Qy	1741	CGGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG	1800
Db	1741	CGGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGCTGAAGAAATCAAG	1800
Qy	1801	CGGAAAGCAGCAAAATGTGAGCAATAAACTGTTGAAGGCTTAAAGATCAAGGGAAGGCTGAA	1860

Db	1801	CGAAAGCGACAGAAATGTTCAGATAAACTGTTGAAGCTTAAGGATCAAGGAAAGCCTGAA	1860
Qy	1861	GTGGGAGATAATGCGMAACTTGGAGAAGATCAATCTCTGAGCAGCAGCTCAAAATTCAGGAG	1920
Db	1861	GTGGGAGATAATGCGAAACTTGGAGAAGATCAATCTCTGAGCAGCAGCTCAAAATTCAGGAG	1920
Qy	1921	CTCCAAGAGAAACTTGGAGAAAGCCT-----	1944
Db	1921	CTCCAAGAGAAACTTGGAGAAAGCCTGTAAAGCCAGCAGCGAGGCCACCGAGCTGCTGCAG	1980
Qy	1945	-----GCAAGGAGCGAGCCGAGAGGGAGCTGGAGAAGCTGGAGAAGCTGCAGAACCCAGAG	1992
Db	1981	AATATTCGCCACAGGCAAAAGGAGCGAGCGAGGGAGCTGGAGAAGCTGCAGAACCCAGAG	2040
Qy	1993	GATTCTTTCTGAAGAGGCATCAGAAAGAGCTGGTGGAAAGCTGAGGAAGCTGAGGAAGCCGCCATTCTCTG	2052
Db	2041	GATTCTTCTGAAGGCATCAGAAAGAGCTGGTGGAAAGCTGAGGAAGCCGCCATTCTCTG	2100
Qy	2053	GAGAACAAAGGTAAAGAGACTTAGAGACCATGGAGCGCTAGAGAAAACAGACTGAAGAGATGAC	2112
Db	2101	GAGAACAAAGGTAAAGAGACTTAGAGACCATGGAGCGCTAGAGAAAACAGACTGAAGAGATGAC	2160
Qy	2113	ATCCAGACAAAATCCCACAAGATCCAGCAGATCGCTGATATAAAATCTTGGAGCTCGGAAGAG	2172
Db	2161	ATCCAGACAAAATCCCACAAGATCCAGCAGATCGCTGATATAAAATCTTGGAGCTCGGAAGAG	2220
Qy	2173	AAACATCGGAGGCCCAAGTCTCAGGCCAGCACTAGAAAGTGCACCTGAAACAGAAAGAG	2232
Db	2221	AAACATCGGAGGCCCAAGTCTCAGGCCAGCACTAGAAAGTGCACCTGAAACAGAAAGAG	2280
Qy	2233	CAGCACTATGAGGAAAGATTAAGTGTTCGACAATTCAGATTAAGAGAAAGACCTGGCTGAC	2292
Db	2281	CAGCACTATGAGGAAAGATTAAGTGTTCGACAATTCAGATTAAGAGAAAGACCTGGCTGAC	2340
Qy	2293	AAGGAGACACTTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAAGGCAAA	2352
Db	2341	AAGGAGACACTTGGAGAACATGATGCAGAGACACGAGGAGGAGGCCCATGAGAAGGCAAA	2400
Qy	2353	ATTCTCAGCGAAACAGAGGGCATGATCAATGCTATCGATTCCAGATCAGATCCCTGGAA	2412
Db	2401	ATTCTCAGCGAAACAGAGGGCATGATCAATGCTATCGATTCCAGATCAGATCCCTGGAA	2460
Qy	2413	CAGAGGATTGTGAACTGTCTGAAGCCAAATAACTTGCAGCAAAATAGCAGTCTTTTACC	2472
Db	2461	CAGAGGATTGTGAACTGTCTGAAGCCAAATAACTTGCAGCAAAATAGCAGTCTTTTACC	2520
Qy	2473	CAAAAGAACATGAAGGCCCAAGAAAGATGATTTCTGAACTCAGGCAACAGAAAATTTTAC	2532
Db	2521	CAAAAGAACATGAAGGCCCAAGAAAGATGATTTCTGAACTCAGGCAACAGAAAATTTTAC	2580
Qy	2533	CTGGAGACACAGGCTGGGAAGTTTGGAGGCCAGAACCCGAAACTTGGAGAGCAGCTGGAG	2592
Db	2581	CTGGAGACACAGGCTGGGAAGTTTGGAGGCCAGAACCCGAAACTTGGAGAGCAGCTGGAG	2640
Qy	2593	AAGATCAGCCACCAAGACCACTGTGACAGATTCGGCTCTGGAATCTGGAGACAGATTG	2652
Db	2641	AAGATCAGCCACCAAGACCACTGTGACAGATTCGGCTCTGGAATCTGGAGACAGATTG	2700
Qy	2653	CGGAGGTCAGTCTAGAGACACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCCACAGAG	2712
Db	2701	CGGAGGTCAGTCTAGAGACACGAGGAGCAGAACTGGAGCTCAAGCGCCAGCTCCACAGAG	2760
Qy	2713	CTACAGCTTCCCTCGAGGAGCGGAGTCAAGTTGACAGCCCTCGAGCTGCACGGCG	2772
Db	2761	CTACAGCTTCCCTCGAGGAGCGGAGTCAAGTTGACAGCCCTCGAGCTGCACGGCG	2820
Qy	2773	GCCTTGAGAGCAGGCTTGCAGGGCGAGACAGAGCTTGGAGAGACACACAGAGCT	2832
Db	2821	GCCTTGAGAGCAGGCTTGCAGGGCGAGACAGAGCTTGGAGAGACACACAGAGCT	2880
Qy	2833	GAAGAGGAGATCAGGCACTCAGGCACATAGAGATGAAATCCAGCGCAAAATTTTATGCT	2892



```
QY 4903 GTGAAGAAAGTGAAACAGTCCCTGGCCAGTCCACACCTGCTGCCAGCCGACATCTCA 4962
Db 5056 GTGAAGAAAGTGAAACAGTCCCTGGCCAGTCCACACCTGCTGCCAGCCGACATCTCA 5115
QY 4963 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGGGGCAGCAAGATTGAAAC 5022
Db 5116 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGGGGCAGCAAGATTGAAAC 5175
QY 5023 GGGCTCTGCATCTGTGAGCAGCATGCCAGCAAAAGTCTCATTTCTCGCTACAAACGAAAC 5082
Db 5176 GGGCTCTGCATCTGTGAGCAGCATGCCAGCAAAAGTCTCATTTCTCGCTACAAACGAAAC 5235
QY 5083 CTGAGCAAAATCTGCATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5142
Db 5236 CTCAGCAAAATCTGCATCCGGAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5295
QY 5143 TTCACCAATTTACAGTATCTCTTGGAAACCAATAAATTTCTACGAAATTCGACATGAAGCAG 5202
Db 5296 TTCACCAATTTACAGTATCTCTTGGAAACCAATAAATTTCTACGAAATTCGACATGAAGCAG 5355
QY 5203 TACAGCTCGAGGAATTCCTGGATAAGATAGACATTCCTTTGGCACCCTGCTGTGTTGCC 5262
Db 5356 TACAGCTCGAGGAATTCCTGGATAAGATAGACATTCCTTTGGCACCCTGCTGTGTTGCC 5415
QY 5263 GCCTCTTCCACAGCTTCCCTGCTCTCAATCGTCAGTGAACGCGCAGGCGAGCAGAG 5322
Db 5416 GCCTCTTCCACAGCTTCCCTGCTCTCAATCGTCAGTGAACGCGCAGGCGAGCAGAG 5475
QY 5323 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTTACGGAAGACGTAGC 5382
Db 5476 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTTACGGAAGACGTAGC 5535
QY 5383 CGACAGACATCTCAAGTGGAGTCTGCTACCTTGGCCCTTGGCCCTACAGAGAACCTAT 5442
Db 5536 CGACAGACATCTCAAGTGGAGTCTGCTACCTTGGCCCTTGGCCCTACAGAGAACCTAT 5595
QY 5443 CTGTTTGTGACCACTTCAACTCACTCGAAGTAAATTGAGATCCAGGACGCTCTCAGCA 5502
Db 5596 CTGTTTGTGACCACTTCAACTCACTCGAAGTAAATTGAGATCCAGGACGCTCTCAGCA 5555
QY 5503 GGAACCCCTCCCGAGGCTACCTGGACATCCCGAACCCCGGCTACCTGGGCCCTGCAAT 5562
Db 5656 GGAACCCCTCCCGAGGCTACCTGGACATCCCGAACCCCGGCTACCTGGGCCCTGCAAT 5715
QY 5563 TCCTCAGAGCGATTACTTGGCGTCTCATACAGATAAATTAAAGGTTCATTGCTGC 5622
Db 5716 TCCTCAGAGCGATTACTTGGCGTCTCATACAGATAAATTAAAGGTTCATTGCTGC 5775
QY 5623 AAGGGAACCTCGTGAAGGAGTCCGGCAGTGAACACACCCGGGGCCCGTCCACCTCCCGC 5682
Db 5776 AAGGGAACCTCGTGAAGGAGTCCGGCAGTGAACACACCCGGGGCCCGTCCACCTCCCGC 5835
QY 5683 AG 5684
Db 5836 AG 5837
```

## RESULT 12

```
US-10-276-774-137
; Sequence 137, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276, 774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
```

```
; SOFTWARE: Custom
; SEQ ID NO 137
; LENGTH: 3131
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-137
```

## Query Match

```
Best Local Similarity 41.6%; Score 2564.8; DB 13; Length 3131;
Matches 2781; Conservative 0; Mismatches 2; Indels 195; Gaps 2;
```

```
QY 2857 GCACATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTGGTAACAGCTGTACTGTATC 2916
Db 1 GCACATAGAGATGAATCCAGCGCAAAATTTGATGCTCTTGGTAACAGCTGTACTGTATC 60
QY 2917 ACAGACCTGGAGGAGCAGCTAAACAGCTGACCGGAGCAACGCTGAACCTCAACAAACAA 2976
Db 61 ACAGACCTGGAGGAGCAGCTAAACAGCTGACCGGAGCAACGCTGAACCTCAACAAACAA 120
QY 2977 AACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCAACGACGAGATTGTACAA 3036
Db 121 AACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCAACGACGAGATTGTACAA 180
QY 3037 CTGCGAAGTGAAGTGGACCATCTCCGCGGAGATCACGGAAACGAGAGATGCAGCTTACC 3096
Db 181 CTGCGAAGTGAAGTGGACCATCTCCGCGGAGATCACGGAAACGAGAGATGCAGCTTACC 240
QY 3097 AGCCAGAAAGCAAAACGATGGAGGCTCTGAAGACCAACGCTGACCATGTGGAGGAACAGGTC 3156
Db 241 AGCCAGAAAGCAAAACGATGGAGGCTCTGAAGACCAACGCTGACCATGTGGAGGAACAGGTC 300
QY 3157 ATGGAATTTGGAGGCCCTTAAACGATGCTGTAGAAAAGAGCGGAGTGGAGGCCCTGG 3216
Db 301 ATGGAATTTGGAGGCCCTTAAACGATGCTGTAGAAAAGAGCGGAGTGGAGGCCCTGG 360
QY 3217 AGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGGTGTGAGGTTCGAGAGCTGCAGAGA 3276
Db 361 AGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGGTGTGAGGTTCGAGAGCTGCAGAGA 420
QY 3277 ATGCTGACACCCGAGAAACAGAGCGGCGAGAGCCGATCAGCGGATCAGCGATCTCGC 3336
Db 421 ATGCTGACACCCGAGAAACAGAGCGGCGAGAGCCGATCAGCGGATCAGCGATCTCGC 480
QY 3337 CAGTGTGTGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTCTCGCTCTGCAGCAGGCT 3396
Db 481 CAGTGTGTGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTCTCGCTCTGCAGCAGGCT 540
QY 3397 CTCAAAAGCAGAAAGCTGAAGGCGGAGAGCCTCTCTGCAAGCTCAATGACCTGGAGAG 3456
Db 541 CTCAAAAGCAGAAAGCTGAAGGCGGAGAGCCTCTCTGCAAGCTCAATGACCTGGAGAG 600
QY 3457 AAGCATGCTATGCTTGAATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGGAACGA 3516
Db 601 AAGCATGCTATGCTTGAATGAATGCCGAGCTTACAGCAGAGCTGGAGACTGGAACGA 660
QY 3517 GAGCTCAAAACAGAGGCTTCTGGAGAGCAAGCCAAATTAACAGCAGAGATGGACCTGCAG 3576
Db 661 GAGCTCAAAACAGAGGCTTCTGGAGAGCAAGCCAAATTAACAGCAGAGATGGACCTGCAG 720
QY 3577 AAAAATCACATTTTCCGCTCTGACCTCAGGAGCTCAGAAAGCTTAGATCGGCTGATCTA 3636
Db 721 AAAAATCACATTTTCCGCTCTGACCTCAGGAGCTCAGAAAGCTTAGATCGGCTGATCTA 780
QY 3637 CTGAAGCAGAGAAAGAGTGAATGGAGTATCAGCTGGAACCAATTCAGGTTCTCTATTCT 3696
Db 781 CTGAAGCAGAGAAAGAGTGAATGGAGTATCAGCTGGAACCAATTCAGGTTCTCTATTCT 840
QY 3697 CATGAAAAGGTGAAAATGGAGGCACTATTCTCAACAAACCAAACTCATTGATTTCTG 3756
Db 841 CATGAAAAGGTGAAAATGGAGGCACTATTCTCAACAAACCAAACTCATTGATTTCTG 900
QY 3757 CAAGCCAAAATGGACCAACCTGCTTAAAAAGAAAAGGGTTTATTTAGTTCACGGAAGAG 3816
Db 3757 CAAGCCAAAATGGACCAACCTGCTTAAAAAGAAAAGGGTTTATTTAGTTCACGGAAGAG 3816
```



APPLICANT: Peyman, John A.  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Ju, Jingfang  
APPLICANT: Li, Li  
APPLICANT: Guo, Xiaojia (Sasha)  
APPLICANT: Patturajan, Meera  
APPLICANT: Spytek, Kimberly A.  
APPLICANT: Eisinger, Shlomit R.  
APPLICANT: Ellerman, Karen  
APPLICANT: Malyankar, Uriel M.  
APPLICANT: Ott, Tatiana  
APPLICANT: Gorman, Linda  
APPLICANT: Zerhusen, Bryan D.  
APPLICANT: Anderson, David W.  
APPLICANT: Zhong, Mei  
APPLICANT: Catterton, Elina  
APPLICANT: Ji, Weizhen  
APPLICANT: Miller, Charles E.  
APPLICANT: Rastelli, Luca  
APPLICANT: Stone, David J.  
APPLICANT: Pena, Carol E. A.  
APPLICANT: Shenoy, Suresh G.  
APPLICANT: Shimkets, Richard A.  
APPLICANT: Rothenberg, Mark E.  
APPLICANT: Leach, Martin D.  
APPLICANT: Agee, Michele L.  
APPLICANT: Barghs, Constance  
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: 21402-482C  
CURRENT APPLICATION NUMBER: US/10/262,511  
CURRENT FILING DATE: 2003-05-28  
PRIOR APPLICATION NUMBER: 60/326,483  
PRIOR FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: 60/373,815  
PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: 60/327,917  
PRIOR FILING DATE: 2001-10-09  
PRIOR APPLICATION NUMBER: 60/381,642  
PRIOR FILING DATE: 2002-05-17  
PRIOR APPLICATION NUMBER: 60/328,029  
PRIOR FILING DATE: 2002-10-09  
PRIOR APPLICATION NUMBER: 60/381,038  
PRIOR FILING DATE: 2002-05-16  
PRIOR APPLICATION NUMBER: 60/328,056  
PRIOR FILING DATE: 2001-10-09  
PRIOR APPLICATION NUMBER: 60/373,260  
PRIOR FILING DATE: 2002-04-17  
PRIOR APPLICATION NUMBER: 60/373,826  
PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: 60/327,435  
PRIOR FILING DATE: 2001-10-05  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 439  
SOFTWARE: CuraSeqList version 0.1  
SEQ ID NO 7  
LENGTH: 2542  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2)..(2542)  
US-10-262-511-7

Query Match 39.2%; Score 2414; DB 13; Length 2542;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 2475; Conservative 5; Indels 48; Gaps 1;

1846 CAAAGGAAGCTCAAGTGGAGAAATATGCGAACTGGAGAGATCAATGCTCAGCAGCAG 1905  
11 CAGGGCAAGCCTGAAGTGGAGAAATATGCGAACTGGAGAGATCAATGCTCAGCAGCAG 70

1906 CTCAAAATTCAGAGCTCCAGAGAACTGGAGAGGCT----- 1944

Db 71 CTCAAAATTCAGAGCTCCAGAGAACTGGAGAGGCTGTAAAAGCCAGCACGGAGGCC 130  
QY 1945 -----GCAAAGGAGCGAGCCGAGAGGAGCTGGAGAAG 1977  
Db 131 ACCGAGCTGCTGCAGAAATATCCGCCAGCAAGGAGCGAGCCGAGAGGAGCTGGAGAAG 190  
QY 1978 CTGCAGAACCCGAGAGGATTTCTTCTGAAGGCATCAGAAAGAGCTGGTGAAGCTGAGGAA 2037  
Db 191 CTGCAGAACCCGAGAGGATTTCTTCTGAAGGCATCAGAAAGAGCTGGTGAAGCTGAGGAA 250  
QY 2038 CGCGCCCATTTCTTGGAGAACCAAGGTAAGAGACTAGAGACCATAGGACCTTAGAGAAAAC 2097  
Db 251 CGCGCCCATTTCTTGGAGAACCAAGGTAAGAGACTAGAGACCATAGGACCTTAGAGAAAAC 310  
QY 2098 AGACTGAAGGATGACATCCAGACAAAATCCACAGAGATCCAGAGATGGCTGATAAAATT 2157  
Db 311 AGACTGAAGGATGACATCCAGACAAAATCCACAGAGATCCAGAGATGGCTGATAAAATT 370  
QY 2158 CTGGAGCTCGAAGAGAAAACATCGGGAGGCCCAAGTCTCAGCCCCAGCACCTTAGAAGTGCAC 2217  
Db 371 CTGGAGCTCGAAGAGAAAACATCGGGAGGCCCAAGTCTCAGCCCCAGCACCTTAGAAGTGCAC 430  
QY 2218 CTGAAAACAGAAAGAGCGACTATGAGGAAAAGATTAAAGTTTGGACAATCAGATAAAG 2277  
Db 431 CTGAAAACAGAAAGAGCGACTATGAGGAAAAGATTAAAGTTTGGACAATCAGATAAAG 490  
QY 2278 AAAGACCTGGCTGACAAAGGAGACACTGGAGAAACATGATGAGAGACACAGAGGAGGAGGCC 2337  
Db 491 AAAGACCTGGCTGACAAAGGAGACACTGGAGAAACATGATGAGAGACACAGAGGAGGAGGCC 550  
QY 2338 CATGAGAGGGCAAAATTTCTCAGCAACAGAAAGCGGATGATCAATGCTATGGATTCCAAG 2397  
Db 551 CATGAGAGGGCAAAATTTCTCAGCAACAGAAAGCGGATGATCAATGCTATGGATTCCAAG 610  
QY 2398 ATCAGATCCCTGGAAACAGAGGATTTGTGAACTGCTCAAGCCCAATAACTTTGAGCAAAAT 2457  
Db 611 ATCAGATCCCTGGAAACAGAGGATTTGTGAACTGCTCAAGCCCAATAACTTTGAGCAAAAT 670  
QY 2458 AGCAGTCTTTTATCCCAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACCTCAGG 2517  
Db 671 AGCAGTCTTTTATCCCAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACCTCAGG 730  
QY 2518 CAACAGAAATTTTACCTGGAGACACAGCTGGGAGTTCAGAGCCCAAGAACTG 2577  
Db 731 CAACAGAAATTTTACCTGGAGACACAGCTGGGAGTTCAGAGCCCAAGAACTG 790  
QY 2578 GAGGAGCAGCTGGAGAGAGATCAGCCACCAAGACACAGTCAAGAAATCGGCTGCTGGAA 2637  
Db 791 GAGGAGCAGCTGGAGAGAGATCAGCCACCAAGACACAGTCAAGAAATCGGCTGCTGGAA 850  
QY 2638 CTGAGACAGATTCGGGAGGTCAGTCTAGAGACAGAGAGAGAACTTGGAGCTCAAG 2697  
Db 851 CTGAGACAGATTCGGGAGGTCAGTCTAGAGACAGAGAGAGAACTTGGAGCTCAAG 910  
QY 2698 CGCCAGCTCACAGAGTACAGCTCTCCCTCAGAGCGCGAGTCAAGTTTGACAGCCCTG 2757  
Db 911 CGCCAGCTCACAGAGTACAGCTCTCCCTCAGAGCGCGAGTCAAGTTTGACAGCCCTG 970  
QY 2758 CAGGCTCAGCGGCGGCTTGGAGAGCAGCTTCGCGAGCGGAGAGCAGAGCTGGAGAG 2817  
Db 971 CAGGCTCAGCGGCGGCTTGGAGAGCAGCTTCGCGAGCGGAGAGCAGAGCTGGAGAG 1030  
QY 2818 ACCACAGCAAGCTGAAGAGGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAG 2877  
Db 1031 ACCACAGCAAGCTGAAGAGGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAG 1090  
QY 2878 CGCAAAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCAAGACCTTGGAGGAGCAGCTA 2937  
Db 1091 CGCAAAATTTGATGCTCTTCGTAACAGCTGTACTGTAAATCAAGACCTTGGAGGAGCAGCTA 1150  
QY 2938 AACAGCTGACCGAGGACAGCTGAATCAACAAACCAAACTTCTACTTGTCCAAACA 2997  
Db 1151 AACAGCTGACCGAGGACAGCTGAATCAACAAACCAAACTTCTACTTGTCCAAACA 1210

QY 2998 CTCGATGAGCTTCGGCCCAACGACGAGATTGTACAACTCGGAGTGAAGTGAACCAT 3057  
DB |||||  
QY 1211 CTCGATGAGCTTCGGCCCAACGACGAGATTGTACAACTCGGAGTGAAGTGAACCAT 1270  
DB |||||  
QY 3058 CTCGCGGGGAGATCACGGAAACGAGAGATCGAGCTTACCAGCAGAGCAAAACGATGGAG 3117  
DB |||||  
QY 1271 CTCGCGGGGAGATCACGGAAACGAGAGATCGAGCTTACCAGCAGAGCAAAACGATGGAG 1330  
DB |||||  
QY 3118 GCTCTGAACACACGTCGACCATCTCTGGAGGACAGCTCATGGATTGGAGGCCCTAAAC 3177  
DB |||||  
QY 1331 GCTCTGAACACACGTCGACCATCTCTGGAGGACAGCTCATGGATTGGAGGCCCTAAAC 1390  
DB |||||  
QY 3178 GATGAGCTGCTAGAAAAGAGCGGCTAGTGGAGGCTCGAGGAGGCTCTGGGTGATGAG 3237  
DB |||||  
QY 1391 GATGAGCTGCTAGAAAAGAGCGGCTAGTGGAGGCTCGAGGAGGCTCTGGGTGATGAG 1450  
DB |||||  
QY 3238 AAATCCAGTTTGAGTGTGGGTTGAGAGCTGCGAGAGAAATGCTGACACCCAGAGAAACAG 3297  
DB |||||  
QY 1451 AAATCCAGTTTGAGTGTGGGTTGAGAGCTGCGAGAGAAATGCTGACACCCAGAGAAACAG 1510  
DB |||||  
QY 3298 AGCAGGCGAGAGCCGATCAGCGGATCACCGAGTCTCGCCAGGTGGTGAGCTGCGAGT 3357  
DB |||||  
QY 1511 AGCAGGCGAGAGCCGATCAGCGGATCACCGAGTCTCGCCAGGTGGTGAGCTGCGAGT 1570  
DB |||||  
QY 3358 AAGGAGCACAGGCTGAGATTCTCGCTCTSCAGCAGGCTCTCAAAGAGCAGAGCTGAAG 3417  
DB |||||  
QY 1571 AAGGAGCACAGGCTGAGATTCTCGCTCTSCAGCAGGCTCTCAAAGAGCAGAGCTGAAG 1630  
DB |||||  
QY 3418 GCGGAGAGCTCTCTACAGCTCANTAGCTGGAGAGAGCATGCTATGCTTGAATG 3477  
DB |||||  
QY 1631 GCGGAGAGCTCTCTACAGCTCANTAGCTGGAGAGAGCATGCTATGCTTGAATG 1690  
DB |||||  
QY 3478 AATGCCGGAAGCTTACAGCAGAGCTGGAGATCTGAACGAGAGCTCAACAGAGGCTTCTG 3537  
DB |||||  
QY 1691 AATGCCGGAAGCTTACAGCAGAGCTGGAGATCTGAACGAGAGCTCAACAGAGGCTTCTG 1750  
DB |||||  
QY 3538 GAAGACAGCCAAATTAACAGCAGAGTGGACCTGCAGAAAAATCAATTTTCGCTG 3597  
DB |||||  
QY 1751 GAAGACAGCCAAATTAACAGCAGAGTGGACCTGCAGAAAAATCAATTTTCGCTG 1810  
DB |||||  
QY 3598 ACTCAGGAGCTGCAAGAGCTTAGATCGGCTGATCTACTGAAGCAGAGAGAGTGAAC 3657  
DB |||||  
QY 1811 ACTCAGGAGCTGCAAGAGCTTAGATCGGCTGATCTACTGAAGCAGAGAGAGTGAAC 1870  
DB |||||  
QY 3658 TTGGAGTATCAGCTGGAACAAATTCAGGTTCTATTCTCATGAAAGGTGAAATGGAA 3717  
DB |||||  
QY 1871 TTGGAGTATCAGCTGGAACAAATTCAGGTTCTATTCTCATGAAAGGTGAAATGGAA 1930  
DB |||||  
QY 3718 GGCATATTCTCAACAGACCAAACTCATGATTTTCTGAGCCCAAAATGACCAACCT 3777  
DB |||||  
QY 1931 GGCATATTCTCAACAGACCAAACTCATGATTTTCTGAGCCCAAAATGACCAACCT 1990  
DB |||||  
QY 3778 GCTAAAAAGAAAGGTTTATTAGTCAGCGAAAGAGAGCCCTGCTTTACCCACACAG 3837  
DB |||||  
QY 1991 GCTAAAAAGAAAGGTTTATTAGTCAGCGAAAGAGAGCCCTGCTTTACCCACACAG 2050  
DB |||||  
QY 3838 GTTCCTCTGAGTAAATGAGCTGAAGTGGCCCTGGAGAGAGAGAAAGCTCGCTGTGA 3897  
DB |||||  
QY 2051 GTTCCTCTGAGTAAATGAGCTGAAGTGGCCCTGGAGAGAGAGAAAGCTCGCTGTGA 2110  
DB |||||  
QY 3898 GAGCTAGAGAAAGCCCTTCAGAGAGCCCGCATCGAGCTCCGCTCCGCGGAGAGAGCT 3957  
DB |||||  
QY 2111 GAGCTAGAGAAAGCCCTTCAGAGAGCCCGCATCGAGCTCCGCTCCGCGGAGAGAGCT 2170  
DB |||||  
QY 3958 GCCACCGCAAGCAACGAGACCAACCAACCCATCCAGCCAGCCACCCGAGGAGAGAG 4017  
DB |||||  
QY 2171 GCCACCGCAAGCAACGAGACCAACCAACCCATCCAGCCAGCCACCCGAGGAGAGAG 2230  
DB |||||  
QY 4018 ATCGCATGTCGCGCATCTGCGGTGCGCAGAGCAGCCAGCCAGTCCATGAGCTGCTG 4077  
DB |||||  
QY 2231 ATCGCATGTCGCGCATCTGCGGTGCGCAGAGCAGCCAGTCCATGAGCTGCTGCTG 2290  
DB |||||

QY 4078 GCCCGGCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCT 4137  
DB |||||  
QY 2291 GCCCGGCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCT 2350  
DB |||||  
QY 4138 AAGAAACGATGACCAACAAATATTCCTCAACGATTCACGATTCACGATTCACGAGCC 4197  
DB |||||  
QY 2351 AAGAAACGATGACCAACAAATATTCCTCAACGATTCACGATTCACGATTCACGAGCC 2410  
DB |||||  
QY 4198 ACAAAGTGTGCTGT 4257  
DB |||||  
QY 2411 ACAAAGTGTGCTGT 2470  
DB |||||  
QY 4258 GAATGTGAGT 4317  
DB |||||  
QY 2471 GAATGTGAGT 2530  
DB |||||  
QY 4318 CCTGCTGA 4325  
DB |||||  
QY 2531 CCTGCTGA 2538  
DB |||||  
RESULT 14  
US-10-262-511-5  
; Sequence 5, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytak, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; CURRENT FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09

```
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 5
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(2497)
; US-10-262-511-5

Query Match      37.6%; Score 2317.2; DB 13; Length 2497;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 2432; Conservative 0; Mismatches 3; Indels 93; Gaps 2;

QY 1846 CAAGGGAGCTGAAGTGGAGAAATATGCGAAACTGGAGAGATCAATCTGAGCGAGAG 1905
DB 11 CAGGGGAGGCTGAAGTGGAGAAATATGCGAAACTGGAGAGATCAATCTGAGCAGCAG 70
QY 1906 CTCAAAATTCAGAGCTCCAGAGAAACTGGAGAGGCT----- 1944
DB 71 CTCAAAATTCAGAGCTCCAGAGAAACTGGAGAGGCTGTAAAGCCAGCAGGAGGCC 130
QY 1945 -----GNAAGGAGCGAGCCGAGAGGGAGCTGGAGAAAG 1977
DB 131 ACAGAGCTGCTGAGAAATATCCCGAGGCAAGAGAGCGAGCGAGAGGAGCTGGAGAA 190
QY 1978 CTCAGAAACCGAGAGATTTCTCTGAAGGCATCAGAAAGAGCTGTGGAAGCTGAGGAA 2037
DB 191 CTCAGAAACCGAGAGATTTCTCTGAAGGCATCAGAAAGAGCTGTGGAAGCTGAGGAA 250
QY 2038 CGCGCGCATTTCTGGAGAAACAAGTAAAGAGACTAGAGACCATGAGAGCTGAGAGAAAC 2097
DB 251 CGCGCGCATTTCTGGAGAAACAAGTAAAGAGACTAGAGACCATGAGAGCTGAGAGAAAC 310
QY 2098 AGACTGAGGATGACATCCAGACAAATCCCAACGATCCAGCAGTCCAGCAGTGTGTAAT 2157
DB 311 AGACTGAGGATGACATCCAGACAAATCCCAACGATCCAGCAGTGTGTAATGATTAAT 370
QY 2158 CTGAGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGCCAGACACCTAGAAAGTGCA 2217
DB 371 CTGAGAGCTCGAAGAGAAACATCGGGAGGCCCAAGTCTCAGCCAGACACCTAGAAAGTGCA 430
QY 2218 CTGAACAGAAAGAGCAGACATATGAGAAAGATTAAGTTGGAACAATCAGATAAG 2277
DB 431 CTGAACAGAAAGAGCAGACATATGAGAAAGATTAAGTTGGAACAATCAGATAAG 490
QY 2278 AAAGACTGTGCTCACAAGAGAGACACTGGAGAACATGATCGAGAGACAGCAGAGGAGGCC 2337
DB 491 AAAGACTGTGCTCACAAGAGAGACACTGGAGAACATGATCGAGAGACAGCAGAGGAGGCC 550
QY 2338 CATGAGAGGGCAAAATTTCTAGCGAACAGAAAGGAGATCAATGCTATGATTTCCAAAG 2397
DB 551 CATGAGAGGGCAAAATTTCTAGCGAACAGAAAGGAGATCAATGCTATGATTTCCAAAG 610
QY 2398 ATCAGATCCCTGGAACAGAGGATTTGGAACTGTCTGAGGCCAAATAAATCTTCAGCAAAAT 2457
DB 611 ATCAGATCCCTGGAACAGAGGATTTGGAACTGTCTGAGGCCAAATAAATCTTCAGCAAAAT 670
QY 2458 AGCAGTCTTTTACCAGAGAACATGAAGGCCCAAGAGAGAGATGATTTCTGAATCTCAGG 2517
DB 671 AGCAGTCTTTTACCAGAGAACATGAAGGCCCAAGAGAGAGATGATTTCTGAATCTCAGG 730
QY 2518 CAACAGAAATTTTACCTGGAGACACAGGCTGGAGGCTTGGAGGCCCAAGAACCTGAC 2577
DB 731 CAACAGAAATTTTACCTGGAGACACAGGCTGGAGGCTTGGAGGCCCAAGAACCTGAC 2637
QY 2578 GAGGAGAGCTGGAGAGAGATCAGCCCAACAGACACAGTGAACAAGATTCGGCTCTCGAA 2637
DB 791 GAGGAGAGCTGGAGAGAGATCAGCCCAACAGACACAGTGAACAAGATTCGGCTCTCGAA 850
QY 2638 CTGAGAGCAAGATTTGCGGGAGGTCAGTCTAGAGACACGAGGAGCAGAACTGGAGCTCAAG 2697
DB 851 CTGAGAGCAAGATTTGCGGGAGGTCAGTCTAGAGACACGAGGAGCAGAACTGGAGCTCAAG 910
QY 2698 CGCAGCTCAAGAGCTACAGCTCTCCCTGAGAGGCGCAGTCAAGTTGACAGCCCTG 2757
DB 911 CGCAGCTCAAGAGCTACAGCTCTCCCTGAGAGGCGCAGTCAAGTTGACAGCCCTG 970
QY 2758 CAGGCTGACCGGGCGGCCCTGGAGAGCAGCTTCGCCAGGCGGAAGACAGAGCTGGAAGAG 2817
DB 971 CAGGCTGACCGGGCGGCCCTGGAGAGCAGCTTCGCCAGGCGGAAGACAGAGCTGGAAGAG 1030
QY 2818 ACCAGAGCAAGCTGAAGAGGAGATCCAGCACTCAGGCACTACGACATAGAGATGAATCCAG 2877
DB 1031 ACCAGAGCAAGCTGAAGAGGAGATCCAGCACTCAGGCACTACGACATAGAGATGAATCCAG 1090
QY 2878 CGCAAAATTTGATGCTCTTCTGTAAGAGCTGACTGTAATCAGACCTGGAGGAGCAGCTA 2937
DB 1091 CGCAAAATTTGATGCTCTTCTGTAAGAGCTGACTGTAATCAGACCTGGAGGAGCAGCTA 1150
QY 2938 AACGAGCTGACCGGAGGACAAAGCTGAACTCAACCAACCAAACTTCTACTTTGTCCAAACA 2997
DB 1151 AACGAGCTGACCGGAGGACAAAGCTGAACTCAACCAACCAAACTTCTACTTTGTCCAAACA 1210
QY 2998 CTCGATGAGGCTTCTGGCGCCACGACAGATGTAACAATCGGAGTGAAGTGGACAT 3057
DB 1211 CTCGATGAGGCTTCTGGCGCCACGACAGATGTAACAATCGGAGTGAAGTGGACAT 1270
QY 3058 CTCGCGCGGAGATCAACGAAACGAGAGATGAGCTTACCAGCCAGAGCAAAACGATGGAG 3117
DB 1271 CTCGCGCGGAGATCAACGAAACGAGAGATGAGCTTACCAGCCAGAGCAAAACGATGGAG 1330
QY 3118 GCTCTGAAGACCAAGTGCACATGCTGGAGAAACAGGTCAATGATTTGGAGGCCCTTAAC 3177
DB 1331 GCTCTGAAGACCAAGTGCACATGCTGGAGAAACAGGTCAATGATTTGGAGGCCCTTAAC 1390
QY 3178 GATGAGCTGCTAGAAAGAGCGCAGTGGAGGCCCTGGAGGAGCGCTCTGGGTGATGAG 3237
DB 1391 GATGAGCTGCTAGAAAGAGCGCAGTGGAGGCCCTGGAGGAGCGCTCTGGGTGATGAG 1450
QY 3238 AAATCCCAGTTTGAAGTCTCGGGTTGAGAGCTCGAGAGAAATGCTGGACACCGAGAAACAG 3297
DB 1451 AAATCCCAGTTTGAAGTCTCGGGTTGAGAGCTCGAGAGAAATGCTGGACACCGAGAAACAG 1510
QY 3298 AGCAGGCGGAGCGGATCAGCGGATCAGCGAGTCTCGCAGGCTGGAGCTGGAGCTGGAGTG 3357
DB 1511 AGCAGGCGGAGCGGATCAGCGGATCAGCGAGTCTCGCAGGCTGGAGCTGGAGCTGGAGTG 1570
QY 3358 AAGAGCACCAAGGCTGAGATTTCTGCTCTCAGCAGGCTCTCAAGAGCAGAAAGCTGAAG 3417
DB 1571 AAGAGCACCAAGGCTGAGATTTCTGCTCTCAGCAGGCTCTCAAGAGCAGAAAGCTGAAG 1630
QY 3418 GCCAGAGCGCTCTGACAAAGCTCAATGACCTGGAGAGAGAGATGCTATGCTGAAATG 3477
DB 1631 GCCAGAGCGCTCTGACAAAGCTCAATGACCTGGAGAGAGAGATGCTATGCTGAAATG 1690
QY 3478 AATCCCCAAGAGCTTACAGCAGAAAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGCTTCTG 3537
DB 1691 AATCCCCAAGAGCTTACAGCAGAAAGCTGGAGACTGAAACGAGAGCTCAAAACAGAGCTTCTG 1750
QY 3538 GAAGAGCAAGCCAAATTTACAGCAGCAGATGGAAGCTGGAGAGAGAGATCAATTTTCCGCTG 3597
DB 1751 GAAGAGCAAGCCAAATTTACAGCAGCAGATGGAAGCTGGAGAGAGAGATCAATTTTCCGCTG 1810
QY 3598 ACTCAGGAGCTGCAAGAGGCTCTAGATCGGGCTGATCTACTGAGAGCAGAAAGAGTGCAC 3657
DB 1811 ACTCAGGAGCTGCAAGAGGCTCTAGATCGGGCTGATCTACTGAGAGCAGAAAGAGTGCAC 1870
```

```
QY 3658 TTGGAGTATCAGTGGAAAAACATTGAGGTTCTTATCTCATGAAAAGGTGAAATGGAA 3717
Db 1871 TTGGAGTATCAGTGGAAAAACATTGAGGTTCTTATCTCATGAAAAGGTGAAATGGAA 1930
QY 3718 GGCATATTTCTCAACAACCAAACTCATTGATTTTCTGCAAGCCAAAATGACCAACCT 3777
Db 1931 GGCATATTTCTCAACAACCAAACTCATTGATTTTCTGCAAGCCAAAATGACCAACCT 1990
QY 3778 GCTAAAAAGAAAAGGGTTTATTTAGTCGACGGAAGAGGACCTTGTCTTTACCCACACAG 3837
Db 1991 GCTAAAAAGAAA-----AG 2005
QY 3838 GTTCTCTGCACTCAATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTTCGTGTGCA 3897
Db 2006 GTTCTCTGCACTCAATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTTCGTGTGCA 2065
QY 3898 GAGCTAGAGGAGCCCTTCAGAAAGCCCGCATCGAGCTCCGGTCCGCCCGGGAGGAAGCT 3957
Db 2066 GAGCTAGAGGAGCCCTTCAGAAAGCCCGCATCGAGCTCCGGTCCGCCCGGGAGGAAGCT 2125
QY 3958 GCCACCGCAAGCAACGAGCCACCAACCCATCCAGCCAGCCACCGGAGGAGGAGGAG 4017
Db 2126 GCCACCGCAAGCAACGAGCCACCAACCCATCCAGCCAGCCACCGGAGGAGGAGGAG 2185
QY 4018 ATCGCATGTCGCGCATCGTGGTCCGAGAGCAGCAGCCAGTGCATGAGCTGCTG 4077
Db 2186 ATCGCATGTCGCGCATCGTGGTCCGAGAGCAGCAGCCAGTGCATGAGCTGCTG 2245
QY 4078 GCCCGGCCATCCAGCCGCAAGAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCT 4137
Db 2246 GCCCGGCCATCCAGCCGCAAGAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCT 2305
QY 4138 AAGGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4197
Db 2306 AAGGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2365
QY 4198 ACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4257
Db 2366 ACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2425
QY 4258 GAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4317
Db 2426 GAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2485
QY 4318 CTTGCTGA 4325
Db 2486 CTTGCTGA 2493
```

## RESULT 15

```
US-10-262-511-3
; Sequence 3, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glendia
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spyttek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
```

```
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 3
; LENGTH: 1870
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1870)
US-10-262-511-3
```

Query Match 29.9%; Score 1841; DB 13; Length 1870;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1844; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 60
Db 14 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATGCTGCTGCTGAACCCATT 73
QY 61 GCCAGCCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120
Db 74 GCCAGCCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 133
QY 121 CAGCAGATGCTCCTCTTTCCGAGAGGATATAGATCCCTCTTTGTTCTTTTGA 180
Db 134 CAGCAGATGCTCCTCTTTCCGAGAGGATATAGATCCCTCTTTGTTCTTTGA 193
QY 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCAGTGAACACTTTGCCGAAGTAT 240
Db 194 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCAGTGAACACTTTGCCGAAGTAT 253
QY 241 TCCGACACCATAGCTGATGATTACAGGAGCTCCAGCTTCGGCAAGGACTTCGAGTCA 300
Db 254 TCCGACACCATAGCTGATGATTACAGGAGCTCCAGCTTCGGCAAGGACTTCGAGTCA 313
QY 301 AGTCTTGTAGGTTGTGGTCACTTTGCTGAAGTGCAGTGGTAAAGAGAGAAAGCAACCGGG 360
Db 314 AGTCTTGTAGGTTGTGGTCACTTTGCTGAAGTGCAGTGGTAAAGAGAGAAAGCAACCGGG 373
```

QY 361 GACATCTATCTATGAAGTGATCAAGAAAGGCTTTATTTGGCCCCAGGAGCAGGTTTCA 420  
DB |||||  
DB 374 GACATCTATCTATGAAGTGATCAAGAAAGGCTTTATTTGGCCCCAGGAGCAGGTTTCA 433  
QY 421 TTTTGTGAGGAAGAGCGGAACAATATTATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 480  
DB |||||  
DB 434 TTTTGTGAGGAAGAGCGGAACAATATTATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 493  
QY 481 CAGTATGCCCTTTTCAGAGCAAAAATACCTTTATCTCATGAGGAAATATCAGCCTGGAGG 540  
DB |||||  
DB 494 CAGTATGCCCTTTTCAGAGCAAAAATACCTTTATCTCATGAGGAAATATCAGCCTGGAGG 553  
QY 541 GACTTCTGTCACCTTTTGAATAGATATGAGGACCACTTAGATGAAAACCTGTATACAGTTT 600  
DB |||||  
DB 554 GACTTCTGTCACCTTTTGAATAGATATGAGGACCACTTAGATGAAAACCTGTATACAGTTT 613  
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCACAGGGTTTCATCTGATGGGATACGTGCATCGA 660  
DB |||||  
DB 614 TACCTAGCTGAGCTGATTTTGGCTGTTTCACAGGGTTTCATCTGATGGGATACGTGCATCGA 673  
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720  
DB |||||  
DB 674 GACATCAAGCCTGAGAACATTTCTCGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 733  
QY 721 GGTCTGCCGCGAATAATGAAATTCMAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780  
DB |||||  
DB 734 GGATCTGCCGCGAATAATGAAATTCMAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 793  
QY 781 CCAGATTACATGCTCTCTGAAGTGCTGACTGTGATGAACGGGATGGAAGGCACCTTAC 840  
DB |||||  
DB 794 CCAGATTACATGCTCTCTGAAGTGCTGACTGTGATGAACGGGATGGAAGGCACCTTAC 853  
QY 841 GGCTCGGACTGTGACTGGTGGTCACTGGCGCGTATTTGCTATGAGATGATTTATGGAGA 900  
DB |||||  
DB 854 GGCTCGGACTGTGACTGGTGGTCACTGGCGCGTATTTGCTATGAGATGATTTATGGAGA 913  
QY 901 TCCCTCTCGCAGAGGAGACCTCTCGCAGAACCTTCAATAACATTTATGAAATTTCCAGGG 960  
DB |||||  
DB 914 TCCCTCTCGCAGAGGAGACCTCTCGCAGAACCTTCAATAACATTTATGAAATTTCCAGGG 973  
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGTATCTGATTCAAAGC 1020  
DB |||||  
DB 974 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGTATCTGATTCAAAGC 1033  
QY 1021 TTGTTGTGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1080  
DB |||||  
DB 1034 TTGTTGTGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTTGAAGTCTTTTGTGCCATCCTTTCTTC 1093  
QY 1081 TCTAAATTTGACTGGAAACAACATTCGTAACTCTCCTCCCTCCCTTCGTTCCACCCCTCAAG 1140  
DB |||||  
DB 1094 TCTAAATTTGACTGGAAACAACATTCGTAACTCTCCTCCCTTCGTTCCACCCCTCAAG 1153  
QY 1141 TCTGACGATGACACCTCCAAATTTTGAATGAACCAAGAGAAATTCGTGGGTTTCATCCTCT 1200  
DB |||||  
DB 1154 TCTGACGATGACACCTCCAAATTTTGAATGAACCAAGAGAAATTCGTGGGTTTCATCCTCT 1213  
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTTCG 1260  
DB |||||  
DB 1214 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTTCG 1273  
QY 1261 TACAGCAAGGCACTGGGATTTCTTGGTATGATCTGATCTGTTGTCTCGGTTCTGGACTCC 1320  
DB |||||  
DB 1274 TACAGCAAGGCACTGGGATTTCTTGGTATGATCTGATCTGTTGTCTCGGTTCTGGACTCC 1333  
QY 1321 CCTGCCAGACTAGCTCCATGGAAAGAACTTCTCATCAAGCAAGCAAGAGCTACAAGAC 1380  
DB |||||  
DB 1334 CTTGCCAGACTAGCTCCATGGAAAGAACTTCTCATCAAGCAAGCAAGAGCTACAAGAC 1393  
QY 1381 TCTCAGGACAGTGTCAAGAATGGAGCAGGAAATGACCCGTTTACATCGGAGAGTGTCA 1440  
DB |||||  
DB 1394 TCTCAGGACAGTGTCAAGAATGGAGCAGGAAATGACCCGTTTACATCGGAGAGTGTCA 1453

QY 1441 GAGGTGGAGGCTGTGCTTAGTTCAGAAAGGCTGAGGCTGAAGGCTCTGAGACTCAGAGA 1500  
DB |||||  
DB 1454 GAGGTGGAGGCTGTGCTTAGTTCAGAAAGGCTGAGGCTGAAGGCTCTGAGACTCAGAGA 1513  
QY 1501 TCCCTCTCTGGAGCAGGACCTTTGCTTACCTACATCAAGAAATGCAAGTAAAGCGAAGT 1560  
DB |||||  
DB 1514 TCCCTCTCTGGAGCAGGACCTTTGCTTACCTACATCAAGAAATGCAAGTAAAGCGAAGT 1573  
QY 1561 TTGGAGCAAGCACCGATGGAGGTGTCACAGAGGATCAAAAGCACTGCAGCTTCTCCAT 1620  
DB |||||  
DB 1574 TTGGAGCAAGCACCGATGGAGGTGTCACAGAGGATCAAAAGCACTGCAGCTTCTCCAT 1633  
QY 1621 GATATCAGAGCAGAGCCCGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGCTCAA 1680  
DB |||||  
DB 1634 GATATCAGAGCAGAGCCCGAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGCTCAA 1693  
QY 1681 GTGGAAGAAATGAGTTTGATGATGAATCAGTTTGAAGAGGATCTTGTCTCAGCAAGAGA 1740  
DB |||||  
DB 1694 GTGGAAGAAATGAGTTTGATGATGAATCAGTTTGAAGAGGATCTTGTCTCAGCAAGAGA 1753  
QY 1741 CGGAGTGTCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATTCAG 1800  
DB |||||  
DB 1754 CGGAGTGTCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATTCAG 1813  
QY 1801 CGGAAAGCGACAGAAATGTCAGCATAAACTGTTGAAGCTTAAGGATCAAG 1849  
DB |||||  
DB 1814 CGGAAAGCGACAGAAATGTCAGCATAAACTGTTGAAGCTTAAGGATCAGG 1862

Search completed: July 3, 2004, 23:37:58

Job time : 1770 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 12:14:48 ; Search time 279 Seconds  
(without alignments)  
12250.690 Million cell updates/sec

Title: US-10-017-216-3  
Perfect score: 6159  
Sequence: 1 atgttgagttcacaatatgg.....ttctgagacacattattgc 6159

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A-COMB.seq:\*
- 2: /cgn2\_6/prodata/2/ina/5B-COMB.seq:\*
- 3: /cgn2\_6/prodata/2/ina/6A-COMB.seq:\*
- 4: /cgn2\_6/prodata/2/ina/6B-COMB.seq:\*
- 5: /cgn2\_6/prodata/2/ina/PCTUS-COMB.seq:\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1391.8	22.6	1515	4	US-09-804-471A-1
2	1391.8	22.6	1515	4	US-10-238-709-1
3	750	12.2	1133	4	US-09-916-204-1
4	258	4.2	258	4	US-09-016-434-513
5	220.8	3.6	2726	2	US-08-422-699A-12
6	220.8	3.6	2726	2	US-08-422-706B-12
7	214	3.5	2511	2	US-08-422-699A-8
8	214	3.5	2511	2	US-08-422-706B-8
9	213.6	3.5	3182	1	US-08-484-044-11
10	205	3.3	174493	4	US-09-804-471A-3
11	205	3.3	174493	4	US-10-238-709-3
12	189.2	3.1	2706	2	US-08-630-822A-61
13	189.2	3.1	2706	2	US-09-005-069-61
14	189.2	3.1	2706	4	US-09-171-158A-20
15	189.2	3.1	2706	4	US-09-004-730A-20
16	189.2	3.1	2706	4	US-08-981-799A-20
17	181.2	2.9	4363	2	US-08-685-576-5
18	179.6	2.9	4848	4	US-09-976-594-295
19	178.8	2.9	48763	4	US-09-916-204-3
20	172.2	2.8	4055	4	US-09-016-434-1105
21	172.2	2.8	4055	4	US-08-685-871-1
22	168.4	2.7	4739	3	US-08-685-576-2
23	150.2	2.4	3323	2	US-08-422-699A-10
24	150.2	2.4	3323	2	US-08-422-706B-10
25	122.6	2.0	3155	4	US-09-442-100-7
26	122.6	2.0	3155	4	US-08-939-106-7
27	122.6	2.0	3155	4	US-09-442-102-7

28	122.2	2.0	1498	4	US-09-509-902A-6	Sequence 6, Appli
29	122.2	2.0	1961	4	US-09-509-902A-15	Sequence 15, Appli
30	122.2	2.0	5276	4	US-09-233-857-2	Sequence 2, Appli
31	117.8	1.9	3018	2	US-08-860-150-6	Sequence 6, Appli
32	117.8	1.9	3018	3	US-09-338-132-6	Sequence 11, Appli
33	111.4	1.8	1935	3	US-08-878-989-11	Sequence 11, Appli
34	111.4	1.8	1935	3	US-09-272-796-11	Sequence 11, Appli
35	109.8	1.8	3213	4	US-09-442-100-5	Sequence 5, Appli
36	109.8	1.8	3213	4	US-08-939-106-5	Sequence 5, Appli
37	109.8	1.8	3213	4	US-09-442-102-5	Sequence 5, Appli
38	105	1.7	3489	2	US-08-728-323A-1	Sequence 1, Appli
39	105	1.7	3489	4	US-09-298-568-1	Sequence 1, Appli
40	105	1.7	3489	4	US-09-410-399-1	Sequence 1, Appli
C 41	105	1.7	32207	2	US-08-770-379-20	Sequence 20, Appli
C 42	105	1.7	32207	3	US-08-757-669A-20	Sequence 20, Appli
C 43	105	1.7	32207	4	US-09-230-71A-20	Sequence 20, Appli
44	104.8	1.7	2160	3	US-09-588-256-1	Sequence 1, Appli
45	101.4	1.6	5720	4	US-09-442-100-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-804-471A-1  
; Sequence 1, Application US/09804471A  
; Patent No. 6479269  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL001164  
; CURRENT APPLICATION NUMBER: US/09/804,471A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1515  
; TYPE: DNA  
; ORGANISM: Human  
US-09-804-471A-1

Query Match 22.6%; Score 1391.8; DB 4; Length 1515;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1396; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1	ATGTTGAAGTTCAATATGGAGCGCGGAATCCCTTTCGATCGTGGTCTGTAACCCATT	60
DB	7	ATGTTGAAGTTCAATATGGAGCGCGGAATCCCTTTCGATCGTGGTCTGTAACCCATT	66
QY	61	GCCAGCGGGGCTCCAGGCTGAATCTCTTCTCCAGGGGAAACACCCCTTATGACTCAA	120
DB	67	GCCAGCGGGGCTCCAGGCTGAATCTCTTCTCCAGGGGAAACACCCCTTATGACTCAA	126
QY	121	CAGCAGATGTCCTCTTCCGAGAGGAGGATATAGATCCCTCTTTGTTCTTTGAA	180
DB	127	CAGCAGATGTCCTCTTCCGAGAGGAGGATATAGATCCCTCTTTGTTCTTTGAA	186
QY	181	GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCACTGAGCACTTTGTCGGGAAGTAT	240
DB	187	GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCACTGAGCACTTTGTCGGGAAGTAT	246
QY	241	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAAGGACTTCGAAGTCAGA	300
DB	247	TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAAGGACTTCGAAGTCAGA	306
QY	301	AGTCTTGTAGTTTGGTCTCACTTCTGAGTCAGGCTGTAAGAGAGAGAACCGGG	360
DB	307	AGTCTTGTAGTTTGGTCTCACTTCTGAGTCAGGCTGTAAGAGAGAGAACCGGG	366
QY	361	GACATCTATGTCATGAAGAGGATGAAGAGGAGGCTTATTTGGCCCGAGGAGGCTTCA	420

Db 367 GACATCTATCTATGAAGTGATGAAGAAGAGGCTTTATTGSCCCAGGAGCGAGTTTCA 426  
QY 421 TTTTGTGAGGAAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Db 427 TTTTGTGAGGAAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 486  
QY 481 CAGTATGCGCTTTTCAGGACAAAATACACCTTTATCTGATGAGGAATATCAGCTGAGGG 540  
Db 487 CAGTATGCGCTTTTCAGGACAAAATACACCTTTATCTGATGAGGAATATCAGCTGAGGG 546  
QY 541 GACTTGTGTCACCTTTGAAATAGATATGAGGACCACTTATAGATGAATAACCTGATACAGTTT 600  
Db 547 GACTTGTGTCACCTTTGAAATAGATATGAGGACCACTTATAGATGAATAACCTGATACAGTTT 606  
QY 601 TACTAGCTGAGCTGATTTGGCTGTTTCAGAGGTTTCTCTGATGGGATACCTGATCGA 660  
Db 607 TACTAGCTGAGCTGATTTGGCTGTTTCAGAGGTTTCTCTGATGGGATACCTGATCGA 666  
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTATT 720  
Db 667 GACATCAAGCCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTATT 726  
QY 721 GGATCTGCGCGGAAATGAATCAACAAGATGGTGAATGCCAACTCCGATTCGGACC 780  
Db 727 GGATCTGCGCGGAAATGAATCAACAAGATGGTGAATGCCAACTCCGATTCGGACC 786  
QY 781 CCAGATTACATGCTCTCTGAAAGTGTGACTGTGATGAACGGGATGGAAGGACCTAC 840  
Db 787 CCAGATTACATGCTCTCTGAAAGTGTGACTGTGATGAACGGGATGGAAGGACCTAC 846  
QY 841 GGCCTGAGCTGACCTGCTGAGTGTGAGTGGCGGTGATTCCTATGAGATGATTTATGGAGA 900  
Db 847 GGCCTGAGCTGACCTGCTGAGTGTGAGTGGCGGTGATTCCTATGAGATGATTTATGGAGA 906  
QY 901 TCCCTCTGCGAGGGAACCTCTGCGAGAACCTTCAATAACATTTATGAATTTCCAGCGG 960  
Db 907 TCCCTCTGCGAGGGAACCTCTGCGAGAACCTTCAATAACATTTATGAATTTCCAGCGG 966  
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTGATCTGATCAAGC 1020  
Db 967 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTGATCTGATCAAGC 1026  
QY 1021 TGTGTTGGCGCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC 1080  
Db 1027 TGTGTTGGCGCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC 1086  
QY 1081 TCTAAAATTTGACTGGAAACAACATTCGTAACTCTCTCTCCCTCTGTTTCCACCTCAAG 1140  
Db 1087 TCTAAAATTTGACTGGAAACAACATTCGTAACTCTCTCTCCCTCTGTTTCCACCTCAAG 1146  
QY 1141 TCTGAGATGACACTCCCAATTTGATGAACAGAGAAATTCGTGGGTTTCATCTCTCT 1200  
Db 1147 TCGAGATGACACTCCCAATTTGATGAACAGAGAAATTCGTGGGTTTCATCTCTCT 1206  
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTCCGCTTTGTGGGGTTTTCG 1260  
Db 1207 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTCCGCTTTGTGGGGTTTTCG 1266  
QY 1261 TACAGCAAGGACTGGGATTTCTGTGATGATCTGAGTCTGTTGTGCTGGGTTGAGCTCC 1320  
Db 1267 TACAGCAAGGACTGGGATTTCTGTGATGATCTGAGTCTGTTGTGCTGGGTTGAGCTCC 1326  
QY 1321 CTTGCCAAGACTAGCTCCATGAAAGAACTTCTCATCAAAAGCAAGAGAGCTACAAGAC 1380  
Db 1327 CTTGCCAAGACTAGCTCCATGAAAGAACTTCTCATCAAAAGCAAGAGAGCTACAAGAC 1386  
QY 1381 TCTCAGGACAAGTGTCAAGAT 1403  
Db 1387 TCTCAGGACAAGTGTCAAGAT 1409

RESULT 2

US-10-238-709-1

; Sequence 1, Application US/10238709  
; Patent No. 6680188  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CLO01164DIV  
; CURRENT APPLICATION NUMBER: US/10/238,709  
; CURRENT FILING DATE: 2002-09-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1515  
; TYPE: DNA  
; ORGANISM: Human  
US-10-238-709-1

Query Match 22.6%; Score 1391.8; DB 4; Length 1515;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1396; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGATGCTGCTGCTGACCCATT 60  
Db 7 ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGATGCTGCTGCTGACCCATT 66  
QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTTCCAGGGGAAACCCCTTTTATGACTCAA 120  
Db 67 GCCAACCGCGGCTCCAGGCTGAATCTGTTTCCAGGGGAAACCCCTTTTATGACTCAA 126  
QY 121 CAGCAGATGCTCTCTTCTTCCGAGAGGATATAGATGCTCTTGTCTCTTTGAA 180  
Db 127 CAGCAGATGCTCTCTTCTTCCGAGAGGATATAGATGCTCTTGTCTCTTTGAA 186  
QY 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCCGGAAGTAT 240  
Db 187 GAATGCACTCAGCTGCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCCGGAAGTAT 246  
QY 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTTGGCAAGGACTTCGAAAGTCAGA 300  
Db 247 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTTGGCAAGGACTTCGAAAGTCAGA 306  
QY 301 AGCTCTGTAGTGTGCTGCTCTTGTGAAAGTCAAGTGTAGAGAGAAAGCAACCGGG 360  
Db 307 AGCTCTGTAGTGTGCTGCTCTTGTGAAAGTCAAGTGTAGAGAGAAAGCAACCGGG 366  
QY 361 GACATCTATCTATGAAGTGAAGAGAGAGGCTTTATTGGCCAGGAGCGAGTTTCA 420  
Db 367 GACATCTATCTATGAAGTGAAGAGAGAGGCTTTATTGGCCAGGAGCGAGTTTCA 426  
QY 421 TTTTGTGAGAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480  
Db 427 TTTTGTGAGAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 486  
QY 481 CAGTATGCTTTTCAGGACAAAATACCTTTTATCTGATGAGGAATATCAGCTGAGGG 540  
Db 487 CAGTATGCTTTTCAGGACAAAATACCTTTTATCTGATGAGGAATATCAGCTGAGGG 546  
QY 541 GACTTGTGTCACCTTTGAAATAGATATGAGGACCACTGATGAATAACCTGATACAGTTT 600  
Db 547 GACTTGTGTCACCTTTGAAATAGATATGAGGACCACTGATGAATAACCTGATACAGTTT 606  
QY 601 TACTAGCTGAGCTGATTTGGCTGTTTCAGAGGTTTCTCTGATGGGATACCTGATCGA 660  
Db 607 TACTAGCTGAGCTGATTTGGCTGTTTCAGAGGTTTCTCTGATGGGATACCTGATCGA 666  
QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTATT 720  
Db 667 GACATCAAGCCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGGTATT 726  
QY 721 GGATCTGCGCGGAAATGAATTCAAACAAGATGGTGAATGCCAACTCCGATTCGGACC 780  
Db 727 GGATCTGCGCGGAAATGAATTCAAACAAGATGGTGAATGCCAACTCCGATTCGGACC 786

```
QY 781 CCAGATTACATGGCTCCCTGAGTCTGACTCTGATGACGGGGATGGAAGGCACTAC 840
Db 787 CCAGATTACATGGCTCCCTGAGTCTGACTCTGATGACGGGGATGGAAGGCACTAC 846
QY 841 GGCCTGGAAGTGTGATGCTGTGCTCAGTGGCGGTGATGCTATGAGATGATTTATGGAGA 900
Db 847 GGCCTGGAAGTGTGATGCTGTGCTCAGTGGCGGTGATGCTATGAGATGATTTATGGAGA 906
QY 901 TCCCCCTTCGAGAGGGAACTCTGCGCAAGACCTTCATACATATGAAATTTCCAGCGG 960
Db 907 TCCCCCTTCGAGAGGGAACTCTGCGCAAGACCTTCATACATATGAAATTTCCAGCGG 966
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGCAGAGTGCATCTTCTTGTATCTGATTCATTAAGC 1020
Db 967 TTTTGAATTTCCAGATGACCCCAAGTGCAGAGTGCATCTTCTTGTATCTGATTCATTAAGC 1026
QY 1021 TTGTTGTCGGCCGAGAGAGAGACTGAAATTTGAAGTCTTTGCTGCCATCCTTCTTC 1080
Db 1027 TTGTTGTCGGCCGAGAGAGAGACTGAAATTTGAAGTCTTTGCTGCCATCCTTCTTC 1086
QY 1081 TCTAAATTTGACTGGAACAACTTCGTAACCTCTCTCCCTCTGTTTCCCACTCAAG 1140
Db 1087 TCTAAATTTGACTGGAACAACTTCGTAACCTCTCTCCCTCTGTTTCCCACTCAAG 1146
QY 1141 TCTGACGATGACCTCCAAATTTTGTATGAACAGAGAAATTCGTGGGTTTCATCTCT 1200
Db 1147 TCCGACGATGACCTCCAAATTTTGTATGAACAGAGAAATTCGTGGGTTTCATCTCT 1206
QY 1201 CGTGCAGCTGAGCCCTCAGGCTTCTCGGTGAGAACTGCCGTTTGTGGGTTTCG 1260
Db 1207 CGTGCAGCTGAGCCCTCAGGCTTCTCGGTGAGAACTGCCGTTTGTGGGTTTCG 1266
QY 1261 TACAGCAAGGCACTGGGATCTTGTGATGATGAGTCTGTTGTGCGGTCTGGACTCC 1320
Db 1267 TACAGCAAGGCACTGGGATCTTGTGATGATGAGTCTGTTGTGCGGTCTGGACTCC 1326
QY 1321 CTTGCCAAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1380
Db 1327 CTTGCCAAGACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1386
QY 1381 TCTCAGGCAAGTGCACAAAGT 1403
Db 1387 TCTCAGGCAAGTGCACAAAGT 1409

RESULT 3
US-09-916-204-1
; Sequence 1, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916, 204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1133
; TYPE: DNA
; ORGANISM: Human
US-09-916-204-1

Query Match 12.2%; Score 750; DB 4; Length 1133;
Best Local Similarity 99.3%; Pred. No. 1.2e-193;
Matches 753; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTTCAAATATGAGCGGGAACTCTTTGGATGCTGCTGCTGGAACCCATT 60
Db 54 ATGTTGAAGTTCAAATATGAGCGGGAACTCTTTGGATGCTGCTGCTGGAACCCATT 113
```

```
QY 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGAAACCCCTTTATGACTCA 120
Db 114 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGAAACCCCTTTATGACTCA 173
QY 121 CACAGATGTCCTCTCTTTCCGAGAAAGGATATTAAGATGCCCTCTTTGTTCTCTTGA 180
Db 174 CACAGATGTCCTCTCTTTCCGAGAAAGGATATTAAGATGCCCTCTTTGTTCTCTTGA 233
QY 181 GAATGCACTCAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAAGTAT 240
Db 234 GAATGCACTCAGCTCTCTGATGAAGATTAAGCAGCTGAGCAACTTTGTCGGAAGTAT 293
QY 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGCGAAAGGACTTCGAAAGTCAGA 300
Db 294 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGCGAAAGGACTTCGAAAGTCAGA 353
QY 301 AGCTTTGTAGTGTGCTGCTGCTGAGTGCAGTGTGTAAGAGAGAAAGCAACCGGG 360
Db 354 AGCTTTGTAGTGTGCTGCTGCTGAGTGCAGTGTGTAAGAGAGAAAGCAACCGGG 413
QY 361 GACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 420
Db 414 GACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTGGCCCGAGGACAGGTTTCA 473
QY 421 TTTTGTGAGAAAGCGGACATATTAATCTCGAAGCAGACGCGCTGGATCCCCCAATTA 480
Db 474 TTTTGTGAGAAAGCGGACATATTAATCTCGAAGCAGACGCGCTGGATCCCCCAATTA 533
QY 481 CAGTATGCTCTTCAGACCAAAATCACTTTATCTGATGAGGAATATCAGCCTGGAGGG 540
Db 534 CAGTATGCTCTTCAGACCAAAATCACTTTATCTGATGAGGAATATCAGCCTGGAGGG 593
QY 541 GACTTGCTGCTATTTGTAATAGATAGAGACCACTGATAGATGAAACCTGATACAGTTT 600
Db 594 GACTTGCTGCTATTTGTAATAGATAGAGACCACTGATAGATGAAACCTGATACAGTTT 653
QY 601 TACTAGCTGAGCTGATTTGGCTGTTCAAGCGTTTCATCTGATGGATACGTCATCGA 660
Db 654 TACTAGCTGAGCTGATTTGGCTGTTCAAGCGTTTCATCTGATGGATACGTCATCGA 713
QY 661 GACATCAAGCTCGAGAACTTCTCGTTGACCGCAGCAGGACACATCAAGCTGGTGGATTT 720
Db 714 GACATCAAGCTCGAGAACTTCTCGTTGACCGCAGCAGGACACATCAAGCTGGTGGATTT 773
QY 721 GGATCTGCCCGGAAATGAATTCAAACAGATGGTGA 758
Db 774 GGATCTGCCCGGAAATGAATTCAAACAGATGGTGA 811

RESULT 4
US-09-016-434-513
; Sequence 513, Application US/09016434
; Patent No. 6509938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PA-0002 US  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 513:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 258 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAINON01  
CLONE: 2290031  
US-09-016-434-513

Query Match 4.2%; Score 258; DB 4; Length 258;

Best Local Similarity 100.0%; Pred. No. 3.4e-60;  
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5051 GCAAAGTCGTCAATTCCTCCGTACAAACGAAACCTCAGCAAAATCTGCATCCGGAAGAGA 5110  
DB 1 GCAAAGTCGTCAATTCCTCCGTACAAACGAAACCTCAGCAAAATCTGCATCCGGAAGAGA 60  
QY 5111 TAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTACCAATACAGTATCCTCATTTGAA 5170  
DB 61 TAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTACCAATACAGTATCCTCATTTGAA 120  
QY 5171 CCAATAAAATTCAGAAATCGACATGAAGCAGTACACGCTCGAGGAATTCCTGGATAAGA 5230  
DB 121 CCAATAAAATTCAGAAATCGACATGAAGCAGTACACGCTCGAGGAATTCCTGGATAAGA 180  
QY 5231 ATGACCATTCCTTGGACCTGCTGTGTGTTTCCGCGCTTTCACCAAGCTTCCTGTCTCAA 5290  
DB 181 ATGACCATTCCTTGGACCTGCTGTGTGTTTCCGCGCTTTCACCAAGCTTCCTGTCTCAA 240  
QY 5291 TCGTGCAGGTGACACGG 5308  
DB 241 TCGTGCAGGTGACACGG 258

## RESULT 5

US-08-422-699A-12  
Sequence 12, Application US/08422699A  
Patent No. 5955265  
GENERAL INFORMATION:  
APPLICANT: Brook, J. David  
APPLICANT: Housman, David B.  
APPLICANT: Shaw, Duncan J.  
APPLICANT: Harley, Helen G.  
APPLICANT: Johnson, Keith J.  
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC  
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02713  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,699A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/422,706  
FILING DATE:  
APPLICATION NUMBER: US 08/023,612  
FILING DATE: 26-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/839,255  
FILING DATE: 20-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/01545  
FILING DATE: 19-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/00253  
FILING DATE: 05-FEB-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB9202485.0  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: MIT-5830A2  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2726 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-422-699A-12

Query Match 3.6%; Score 220.8; DB 2; Length 2726;  
Best Local Similarity 53.7%; Pred. No. 2e-49;  
Matches 531; Conservative 0; Mismatches 442; Indels 15; Gaps 3;  
QY 194 CTGCTCATGATGAAGATTAAAGCAGCTGAGCAACTTGTCCGGAAGTATTCGACACCATAG 253  
DB 243 CCGAATCGGCCAGGACAGTAGTCTGCGCCACTTCTTGCACTGGCGGAGCCCATGTTG 302  
QY 254 CTGAGTTACAGGAGCTCCAGCCTTCGGCAAGAGACTTCGAAAGTCAGAAAGTCTTTGTTAG 313  
DB 303 TGAGGCTTAAAGAGGTCCGACTCGAGAGGACGACTTCGAGATTCTGAAGTGTATCGGAC 362  
QY 314 GTGTCACATTTGCTGAAGTGCAGGTGGTAAGAGAGAAAGCAACCGGGGACATCTATGCTA 373  
DB 363 CGCGGGCTTCAGCAGGAGTAGCGGTAGTAGAGTGAAGCAGAGCGGCCAGGTATGCCA 422  
QY 374 TGAAGTGTATGAAGAAAGAGCTTTTATTCGCCAGGAGGAGTTTCATTTTGTAGGAAG 433  
DB 423 TGAAGATCATGAACAAGTGGACATCTGTAAGAGGGGCGAGGTGTCTGCTTCCGTGAGG 482  
QY 434 AGCGGAACATATTATCTCGNAGCACAAGCCCGTGGATCCCCCAATTACAGTATGCTTTTC 493  
DB 483 AGAGGAGCTGTTGGTGAATGGGACCGCGGTGGATCAGCAGCTGACACTTCGCGCTTCC 542  
QY 494 AGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCCTCGAGGGGAGCTTGTGTCTAC 553  
DB 543 AGGATGAGAACTACCTGTACCTGGTTCATGGAGTATTACCTGGCGGGGACCTGCTGACAC 602  
QY 554 TTTTGAATAGATAGAGACCAAGTATGATAAGAAACCTGATACAGTTTACCTAGCTGAGC 613  
DB 603 TGCTGAGCAAGTTTGGGAGCGGATTCGCGCGAGATGCGCGCTTCTACCTGGGGAGA 662  
QY 614 TGATTTTGGCTGTTACAGCGTTTCATCTGTGAGGATACGTGATCGAGACATCAAGCCTG 673

```

663 TTGTCATGGCCATAGACTCGGTGCACCGGCTTGGCTACGTGACAGGGACATCAAAACCG 722
QY
674 AGAACATCTCTGTTGACCGCACAGGACACATCAAGCTGTGGTGGATTTGGATCTGGCGGA 733
Db
723 ACAACATCTCTGTTGACCGCACAGGACACATCAAGCTGTGGTGGATTTGGATCTGGCGGA 782
QY
734 AAATGAATTCACAAAGATGGTGAATGCCAAATCCCGATTGGGACCCAGATTTACATGG 793
Db
783 AGCTGCGGGCAGATGGAACGGTGGCTGGTGGCTGTGGGACACCCAGACTACCTGT 842
QY
794 CTCCTGAAGTCTGA----CTGTGATGAACGGGATGAAAGGACCTACCGCTGGACT 850
Db
843 CCCCCGAGATCCTGACGCTGTGGGCGGTGGGCGGTGGGACAGCAGCTACGGGCGCGAGT 902
QY
851 GTGACTGGTGTGATGCTGGGCGGTGATTGCTCTATCAGATGATTTATGGAGATCCCGCTTCG 910
Db
903 GTGACTGGTGGCGCTGGTGTGATTGCTCTATGAAATGTTCTATGGGACAGCCCTTCT 962
QY
911 CAGAGGAACTCTGCGCAGAACCTTCAATAACATTAATGAATTCAGCGGTTTTGAAAT 970
Db
963 ACGGGATTCACGGCGGAGACCTATGGCAAGATCGTCCACTACAGGAGCACCTCTCTC 1022
QY
971 TTCCAGATGACCCCAAAAGTGAAGT---GACTTTCTGATCTGATTCATCAAAAGCTTGTGT 1027
Db
1023 TGCGCTGGTGGAGGAAGGGTCCCTGAGGAGGCTCGAGCTTCATTCAGCGTGTGTGT 1082
QY
1028 CGGCCAGAAAGAGAGACTGAAATTTGAAGTCTTTGCTGCT-----CATCTTTCT 1078
Db
1083 GTCCCCCGGACACAGCGCTGGCGGGGTGGAGAGCGACTTCCGACACATCCCTTCT 1142
QY
1079 TCTCTAAATTTGACTGGAACAAACATTCGTAACCTCTCTCCCGCTTCTGTTCCACCCCTCA 1138
Db
1143 TCTTGGCTCGACTGGATGGTCTCGGGACAGCGTGGCGGCTTACACCGGATTTGG 1202
QY
1139 AGTCTGACATGACACTCCAAATTTGA 1166
Db
1203 AAGTGGCCACCGACACATGCAACTTGA 1230

```

RESULT 6

```

US-08-422-706B-12
; Sequence 12, Application US/08422706B
; Patent No. 5977333
; GENERAL INFORMATION:
; APPLICANT: Brook, J. David
; APPLICANT: Housman, David E.
; APPLICANT: Shaw, Duncan J.
; APPLICANT: Harley, Helen G.
; APPLICANT: Johnson, Keith J.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
; DYSTROPHY GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02713
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,706B
; FILING DATE: 14-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,543
; FILING DATE: 08-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/023,612

```

```

; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/839,255
; FILING DATE: 20-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01545
; FILING DATE: 19-FEB-1993
; PRIOR APPLICATION DATA: PCT/GB93/00253
; FILING DATE: 05-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB9202485.0
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-5830A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-422-706B-12

Query Match      3.6%; Score 220.8; DB 2; Length 2726;
Best Local Similarity 53.7%; Pred. No. 2e-49;
Matches 531; Conservative 0; Mismatches 442; Indels 15; Gaps 3;

QY 194 CTCCTGATGAAGATTAAAGCAGTGAAGCACTTTGTCGGGAAGTATTCGACACCATAG 253
Db 243 CCGAATCGGCCAGGACAAAGTACGTGGCGGACTTCTTGCACTGGGCGGACCATCGTG 302
QY 254 CTGAGTTACAGGAGCTCCAGCCCTTCGGCAAGACTTCGAACTTCAGAGTCTTTAGGTT 313
Db 303 TGAGGCTTAAGGAGTCCGACTGCAGAGGACGACTTCGAGATTCTGAAGTATCGGAC 362
QY 314 GTGGTCACTTTGCTGAAGTGCAGGTGTGTAAGAGAGAAAGCAACCGGGAGACATCTATG 373
Db 363 CGCGGGCGTTTCAAGGAGTGTAGCGGTAGTGAAGATGAAGCAGACGGCGGAGGTATGCCA 422
QY 374 TGAAGTGTGAAGAAAGAGGCTTTATTGGCCAGGAGCAGGTTTCATTTTTGAGGAAG 433
Db 423 TGAAGTATGAACAAGTGGGACATGCTGAAGAGGGCGAGGTGTGCTTCCTGAGG 482
QY 434 AGCGAACATATTATCTCGAAGCAAGCCCGTGGATCCCAATTAAGATAGCTTTC 493
Db 483 AGAGGACCGTGTGGTGAATGGGACCGCGGTGGATCAGCAGCTGCACATTCGCTTCC 542
QY 494 AGGACAAAATCACCTTTATCTGATGGAGGAATATCAGCTCGAGGGGACTTGTGTAC 553
Db 543 AGGATGAGAACTACCTGTACCTGGTATGAGTATTTACGTGGCGGGGACCTGCTGAC 602
QY 554 TTTTGAATAGATATGAGGACCAAGTTAGATGAAAACTGATACAGTTTACCTAGCTGAGC 613
Db 603 TGCTGAGCAAGTTTGGGAGCGGATTCGCGCGAGATGGCGGCTTCTACCTGGCGGAGA 662
QY 614 TGATTTTGGCTGTTCACAGCGTTTCATCTGATGGGATAGTGCATCGAGACATCAAGCTG 673
Db 663 TTGCTATGGCCATAGACTCGGTGTCACCGCTTGGCTAGTGCACAGGAGACATCAACCCG 722
QY 674 AGAACATTTCTGTTGACCCGACAGACACATCAAGCTGGTGGATTTTGGATCTGCGCGGA 733
Db 723 ACAACATCTCTGTTGAGCCCTGTGGCCACATCCGCTGGCGGACTTCGCTTCTTGCCTCA 782
QY 734 AAATGAATTCACAAAGATGGTGAATGCCAACTCCCGATTGGGACCCAGATTTACATGG 793
Db 783 AGCTGGGCGAGATGGAACGGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 842
QY 794 CTCCTGAAGTGTGCA---CTGTGATGAACGGGAGTGGAAAAAGGACACCTACGCGCTTGGACT 850

```

Db 843 CCCCCGAGATCTCTCAGGCTGTGGCGGTGGGCTGGGACAGGCGAGCTACGGGCCCGAGT 902  
Qy 851 GTGACTGGTGTGATGAGTGGGCTGATTCCTCTATGAGATGATTTATGGGAGATCCCCCTTCG 910  
Db 903 GTGACTGGTGGGCGCTGGGTGATTCCTCTATGAGATGATTTATGGGAGATCCCCCTTC 962  
Qy 911 CAGAGGAACTCTGCCAGAACCTTCAATAACATTTATGAATTTCCAGCGGTTTTCGAAT 970  
Db 963 ACSCGGATTCACGGCGGAGACCTATGGCAAGATCGTCCACTACAAGGAGCACCTCTCTC 1022  
Qy 971 TTCAGATGACCCCAAGTGGAGT---GACATTTCTTGTATGATTCAGATTTGTTGT 1027  
Db 1023 TGCGCTGGTGGAGAGGGGTCCCTGAGGAGGCTCGAGACTTCATTCAGCGGTTGCTGT 1082  
Qy 1028 GCGGCCAGAGAGAGACTGAAGTTTGAAGTCTTTGCTGC-----CATCTTTCT 1078  
Db 1083 GTCCCCCGGAGACACGCTGGCGGGGTGGAGCAGCGGACTTCGGACACATCCCTTCT 1142  
Qy 1079 TCTCTAAATGACTGGAACAACTTCGTAACTCTCTCTCCCTTCCTTCGTTCCACCTCA 1138  
Db 1143 TCTTTGGCTCGACTGGAGTGTCTCCGGACAGCGTGGCCCCCTTTACACCGGATTCG 1202  
Qy 1139 AGTCTGACGATGACACTCCAAATTTGA 1166  
Db 1203 AAGGTGCCACGACACATGCACTTGA 1230

## RESULT 7

US-08-422-699A-8  
; Sequence 8, Application US/08422699A  
; Patent No. 5985265  
; GENERAL INFORMATION:  
; APPLICANT: Brook, J. David  
; APPLICANT: Housman, David E.  
; APPLICANT: Shaw, Duncan J.  
; APPLICANT: Harley, Helen G.  
; APPLICANT: Johnson, Keith J.  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC  
; TITLE OF INVENTION: DYSPROPHY GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02713  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/422,699A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/422,706  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/023,612  
; FILING DATE: 26-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/839,255  
; FILING DATE: 20-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/01545  
; FILING DATE: 19-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB93/00253  
; FILING DATE: 05-FEB-1993  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB9202485.0  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: MIT-5830A2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2511 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1746  
; US-08-422-699A-8  
  
Query Match 3.5%; Score 214; DB 2; Length 2511;  
Best Local Similarity 54.2%; Pred. No. 1.4e-47;  
Matches 508; Conservative 0; Mismatches 415; Indels 15; Gaps 3;  
  
Qy 244 GACACATAGCTGAGTTACAGGAGCTCCAGCCTTCGSCAAAGACTTCGAGTTCAGAGT 303  
Db 43 GCCATCGTGTGAGGCTTAAGGAGGTCGACTGCAGAGGACGACTTCGAGATTCGAAG 102  
Qy 304 CTTGTAGTTGTGTCATTTCTGAGTGCAGTGTGTAAGAGAGAAAGCAACCGGGAC 363  
Db 103 GTGATCGACCGGGCGGTTACGCGAGGTAGCGGTAGTGAAGATGAAGCAGACGCGGCCAG 162  
Qy 364 ATCTATGCTATGAAAGTGTATGAAAGAGAGGCTTTATTGGCCCGAGGAGAGGTTTCATTT 423  
Db 163 GTGTATGCCATGAAGATCATGAACAAGTGGACATCTGAAGAGGGCGGAGGTGTCGTGC 222  
Qy 424 TTTGAGGAGAGCGGACATATTTATCTGAGACACAAGCCCGTGGATCCCCCAATTACAG 483  
Db 223 TTCGCTGAGGAGGAGGACGCTGTTGGTGAATGGGACCGCGCGTGGATCAGCAGCTGCAC 282  
Qy 484 TATGCTCTTTCAGGACAAATAATCACCTTTATCTGATGGAGAAATATCAGCCTTGGAGGGAC 543  
Db 283 TTCGCTTTCAGGATGAGAACTACCTGTACCTGGTCTATGGATTTACGTTGGCGGGGAC 342  
Qy 544 TTGCTGTACTTTTGAATAGATATGAGGACCAAGTTAGATGAAACCTGATACAGTTTAC 603  
Db 343 CTGCTGACACTGCTGAGCAAGTTTGGGAGCGGATTCGCGCGAGATGCGGCTTCTAC 402  
Qy 604 CTAGCTGAGCTGATTTGGCTGTACACGCTTCTATGATGGATACGTCATCGAGAC 663  
Db 403 CTGCGGAGATTTCTATGCCCATAGACTCGGTGACCCGCTTGGCTTACGTCGACAGGAC 462  
Qy 664 ATCAAGCTTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGATTTTGA 723  
Db 463 ATCAAAACCGGACAAACATCTCTGTGGACCGCTGTGGCCACATCCGCTTGGCCGCTTCCG 522  
Qy 724 TCTGCGCGAAATGAATTCACAAAGATGTTGAATCCAACTCCCGATTGGGACCCCA 783  
Db 523 TCTTGCTCAAGCTGCGGCGAGATGAAACGTTGCGGTGCTGCTGGTGTGGGACCCCA 582  
Qy 784 GATTACATGGCTCTTGA---AGTGTGACTGTGATGAACGGGGATGGAAAGGCACTTAC 840  
Db 583 GACTACCTGTCCCCGAGATCCTGACGCTGTGGCGGTGGGCGCTGGGACAGCAGTAC 642  
Qy 841 GGCCTGGAATGACTGTGGTGTGCTGAGTGGGCTGATGCTATGATGATGATTTATGGGAGA 900  
Db 643 GGGCCGAGTGTGACTGTGGTGGCGCTGATTCGCTTATGAAATGTTCTATGGGAG 702  
Qy 901 TCCCCCTTCGAGAGGGAACCTTCGCCAGAACTTCATATAACATTTATGAATTTCCAGCGG 960  
Db 703 ACGCCTTCTACGCGGATTCACGCGGAGACCTTATGCAAGATCGTCCACTACAGAG 762

QY 961 TTTTGAATTCAGATGACCCCAAGTGAAGT---GACTTTCTTGTATCTGATTCAA 1017  
Db 763 CACTCTCTCTGCGCGTGGTGGAGAGGGTCCCTGAGGAGCTCGAGACTTCATTAC 822  
QY 1018 AGCTTGTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGC----- 1068  
Db 823 CGGTGTGTGTCCTCCCGAGACACGCTGCGCGGGTGGAGCAGGACTTCGGACA 882  
QY 1069 CATCTTTCTCTCTAAATTTGACTGGAACAACATTGCTGTAATCTCTCTCCCTTCGTT 1128  
Db 883 CATCTCTCTCTTTGCGCTCGACTGGATGGTCTCGGAGACAGCTGCCCCCTTTACA 942  
QY 1129 CCCACCTCAAGTCTGACGATGACACCTCCAAATTTGA 1166  
Db 943 CCGAATTCGAAGTGGCCACGACACATGCACTTGA 980

## RESULT 8

US-08-422-706B-8  
; Sequence 8, Application US/08422706B  
; Patent No. 5977333  
; GENERAL INFORMATION:  
; APPLICANT: Brook, J. David  
; APPLICANT: Housman, David E.  
; APPLICANT: Shaw, Duncan J.  
; APPLICANT: Harley, Helen G.  
; APPLICANT: Johnson, Keith J.  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC  
; TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/422,706B  
; FILING DATE: 14-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/284,543  
; FILING DATE: 08-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/023,612  
; FILING DATE: 26-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/839,255  
; FILING DATE: 20-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/01545  
; FILING DATE: 19-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB93/00253  
; FILING DATE: 05-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB9202485.0  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: MIT-5830A2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 617-861-9540  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 2511 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1746  
; US-08-422-706B-8

Query Match  
Best Local Similarity 3.5%; Score 214; DB 2; Length 2511;  
Best Local Similarity 54.2%; Pred. No. 1.4e-47;  
Matches 508; Conservative 0; Mismatches 415; Indels 15; Gaps 3;

QY 244 GACACCATAGCTGAGTTACAGGAGCTCCAGCCCTTGGCAAGAGACTTCCAAAGTCAGAAGT 303  
Db 43 GCCATCGTGTGAGGCTTAAAGGAGGTCGAGAGGACGACTTCGAGATTCTGAAG 102  
QY 304 CTTGTAGGTTGTGTCACATTTTGTGAAGTGCAGAGTGTAAAGAGAGAAAGCAACCCGGGAC 363  
Db 103 GTGATCGGACGCGGGCGGTTACGAGAGTACGGTAGTGAAGATGAAGCAGACGGGCCAG 162  
QY 364 ATCTATCTATGAAGTGTATGAAGAAGAGGCTTTATTTGGCCCCAGGAGCAGGTTTCATT 423  
Db 163 GTGTATGCCATGAAGATCATGAACAAGTGGGACATGCTGAAGAGGGGCGAGGTGTCGTG 222  
QY 424 TTTGAGGAGAGCGGAAACATATTATCTCAAGCACAAAGCCCGTGGATCCCCCAATTACAG 483  
Db 223 TTCCGTGAGGAGGAGCGTGTGTTGTTGATGGGACCGCGGTGGATCACGACGCTGCAC 282  
QY 484 TATGCCCTTCAGGACAAAATACCTTTATCTGTATGGAGGAATATCAGCCTCGAGGGGAC 543  
Db 283 TTCGCCCTTCAGGATGAGAACTACCTGTACTCTGTTGATGGAGTATTAAGTGGGCGGG 342  
QY 544 TTGCTGTCACTTTTGAATAGATATGAGGACCAAGTTAGATGAAGAACTGTATACAGTTTAC 603  
Db 343 CTGCTGACACTGCTGAGCAAGTTTGGGAGCGGATTCCGGCCGAGATGGCGCGCTTCTAC 402  
QY 604 CTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTTCATCTGTATGGGATACGTGATCGAGAC 663  
Db 403 CTGGCGGAGATTGTCTATGCCCATAGACTCGGTGCAACCGCTTGGCTTGGCTGGCGGAC 462  
QY 664 ATCAAGCCTGAGAACTTCTGTTGACCCACAGGACACATCAAGCTGGTGGATTGGA 723  
Db 463 ATCAACCCGACACATCTGCTGGACCGCTGTGGCCACATCCGCTGGCGACTTCGGC 522  
QY 724 TCTGCCCGGAAATGAATTCAAACAGATGTTGAATGCAAACTCCGATTTGGGACCCCA 783  
Db 523 TCTTGCTCAAGCTGCGGCGAGATGGAACGCTGCGGTGCTGCTGGTGGTGGCCACCCCA 582  
QY 784 GATTACATGGCTCCTCA---AGTGTGCTGATGATGAACGGGATGGAAGGACCTTAC 840  
Db 583 GACTACTGTCCTCCCGAGATCTGAGGCTGTGGCGGTGGGCTGGGACAGGAGTAC 642  
QY 841 GGCTGGACTGTGACTGGTGGTCAAGTGGCGGTGATGGCTATGAGATGATTTATGGGAGA 900  
Db 643 GGGCCCGAGTGTGACTGTGGCGGTGGTGTATTCGCTATGAATTTCTATGGGAG 702  
QY 901 TCCCCCTTCGACAGGGAACCTCTGCGAAGCTTCAATAACATTAATGAATTTCCAGCGG 960  
Db 703 ACGCCCTTCTACCGGATTCACGCGGAGACATATGCAAGATCTCTCACTACAGAGAG 762  
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGT---GACTTTCTTGTATCTGATTCAA 1017  
Db 763 CACTCTCTCTGCGCTGGTGGAGCAAGGGTCCCTGAGGAGGCTCGAGACTTCATTAC 822  
QY 1018 AGCTTGTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGC----- 1068  
Db 823 CGGTGTGTGTCCTCCCGAGACACGCTGCGCGGGTGGAGCAGGACTTCGGACA 882  
QY 1069 CATCTTTCTCTCTAAATTTGACTGGAACAACATTGCTGTAATCTCTCTCCCTTCGTT 1128  
Db 883 CATCTCTCTCTTTGCGCTCGACTGGATGGTCTCGGAGACAGCTGCCCCCTTTACA 942

QY 1129 CCCACCTCAAGTGTGAGATGACACCTCCAAATTTGA 1166  
 Db 943 CCGGATTCGAAGGTGCGCCGACACATGCAACTTCGA 980

## RESULT 9

US-08-484-044-11  
 ; Sequence 11, Application US/08484044  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Caskey, C. T.  
 ; APPLICANT: Fu, Ying-Hui  
 ; APPLICANT: Friedman, David L.  
 ; APPLICANT: Pizzuti, Antonio  
 ; APPLICANT: Fenwick, Raymond G.  
 ; TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESS: Fulbright & Jaworski, L.L.P.  
 ; STREET: 1301 McKinney, Suite 5100  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: U.S.A.  
 ; ZIP: 77010-3095  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patencin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/484,044  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/019,940  
 ; FILING DATE: 19-FEB-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Paul, Thomas D.  
 ; REGISTRATION NUMBER: 32,714  
 ; REFERENCE/DOCKET NUMBER: D-5443  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 713/651-5325  
 ; TELEFAX: 713/651-5246  
 ; TELEX: 762829  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3182 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-08-484-044-11

Query Match 3 5%; Score 213.6; DB 1; Length 3182;  
 Best Local Similarity 54.2%; Pred. No. 2e-47;  
 Matches 507; Conservative 0; Mismatches 414; Indels 15; Gaps 3;  
 QY 246 CACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAGGACTTCGAAATCAGAAGTCT 305  
 Db 743 CATCGTGGTGAAGCTTAAGGAGTCCGACTGCAGAGGAGGACACTTCGAGATTCGAAGGT 802  
 QY 306 TGTAGTTGTGGTCACTTGTGTAAGTCCAGTGGTGAAGAGAAAGCAACCGGGGACAT 365  
 Db 803 GATCGGACGGGGGGGTTTCAGCGAGGTAGCGGTAGTGAAGATGAAGCAGACGGGCCAGGT 862  
 QY 366 CTATGCTATGAAGTGAAGAAAGAGGCTTTATTGGCCCGCAGGAGCGAGTTTCATTTT 425  
 Db 863 GTATGCCATGAAGATCATGAACAAGTGGACATCTGAAAGAGGGCGGAGGTGTCGCTT 922  
 QY 426 TGAGAGAGCGGACACATATTATCTGAAGACAAAGCCCGTGGATCCCCCAATTTACAGTA 485  
 Db 923 CCGTGAGGAGGAGGACGCGTGTGGTGAATGGGGACCGCGGTGGATCATCAGCAGTGCACCT 982

QY 486 TGCCTTTAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCTTCGGAGGGACTT 545  
 Db 983 CGCCTTCAGGATGAGAACTACCTGTACTTGGTCAATGGAGTATTACGTGGGGGGGACCT 1042  
 QY 546 GCTGTCACTTTGAATAGATATGAGGACCAAGTATAGATGAAAAACCTGATACAGTTTACCT 605  
 Db 1043 GCTGACACTGTGTAGCAAGTTTGGGGAGCGGATTCCGGCCGAGATGGCGCTTCTACCT 1102  
 QY 606 AGCTGAGCTGATTTTGGCTGTTCACAGGTTTCATCTGATGGGATACGTGATCAGACAT 665  
 Db 1103 GCGGAGATTGTCTAGGCGCATAGACTCGGTGCACCGGCTTGGCTACGTTGACAGGGACAT 1162  
 QY 666 CAAGCTCAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTTGGATC 725  
 Db 1163 CAACCCGACAAATCTCTGTGGACCGCTGTGGCCACATCCGCTGGCCGACTTCGGCTC 1222  
 QY 726 TGCCGCGAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACCCAGA 785  
 Db 1223 TTGCTCAAGCTGCGGGCAGATGGAACGTCGCTGCTGGTGGCTGTGGGCACTCCAGA 1282  
 QY 786 TTACATGGCTCCTGA---AGTGTGACTGTGATCAACGGGATGGAAGGACCTACCG 842  
 Db 1283 CTACTGTCCCCGAGATCTCTGAGGCTGTGGGCGGTGGGCTGGGACAGGACGCTACGG 1342  
 QY 843 CTTGACTGTGACTGGTGGTCAATGGGCGGATTTGCTATGAGATGATTTATGGGAGATC 902  
 Db 1343 GCGGAGTGTGACTGTGGGCGCTGGGTGATTTCGCTATGAAATGTTCTATGGGACAG 1402  
 QY 903 CCCCTTCGAGAGGAACTCTCTGCAGAACTTCAATAACATTATGATTTCCAGCGGTT 962  
 Db 1403 GCGCTTCTACGCGATTCCAGCGGAGACATATGGCAAGATCGTCCACTACAGGAGCA 1462  
 QY 963 TTTGAAATTTCCAGATGACCCCAAGTGAGCAGT---GACTTTTCTGATCTGATTTCAAAG 1019  
 Db 1463 CCTCTCTCTGCGCTGGTGACGAAGGGTCCCTGAGGAGGCTCGAGACTTCATTACGG 1522  
 QY 1020 CTTGTTGTGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGC-----CA 1070  
 Db 1523 GTTGTGTGTCCCGGAGACAGCGCTGGGCGGGTGGAGCAGGAGCTTCGGGACACA 1582  
 QY 1071 TCCTTTCTCTCTAAATTTGACTGGAACAACATTCGTAATCTCTCTCCCTCCCTTCGTTCC 1130  
 Db 1583 TCCTTCTCTCTTTGGCTGACTGGGATGCTCTCCGGGACAGCGTGCCTTTTACAC 1642  
 QY 1131 CACCTCAAGCTGACGATGACACCTCCAAATTTGA 1166  
 Db 1643 GGATTTGGAAGGTGCCACCGACACATGCAACTTCGA 1678

## RESULT 10

US-09-804-471A-3  
 ; Sequence 3, Application US/09804471A  
 ; Patent No. 6479269  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WEBSTER, Marion et al  
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
 ; TITLE OF INVENTION: THEREOF  
 ; FILE REFERENCE: CL001164  
 ; CURRENT APPLICATION NUMBER: US/09/804,471A  
 ; CURRENT FILING DATE: 2001-03-13  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 174493  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(174493)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; US-09-804-471A-3

Query Match 3.3%; Score 205; DB 4; Length 174493;  
Best Local Similarity 100.0%; Pred. No. 5.7e-44;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 GGTGAATGCCAAACTCCCGATTGGGACCCAGATTACATGCTCTGAAGTCTGACTGT 812  
Db 130289 GGTGAATGCCAAACTCCCGATTGGGACCCAGATTACATGCTCTGAAGTCTGACTGT 130348

QY 813 GATGAACGGGGATGAAAGGACCTAGCGCTGGACTGTGACTGGTGGTGGGGCGT 872  
Db 130349 GATGAACGGGGATGAAAGGACCTAGCGCTGGACTGTGACTGGTGGTGGGGCGT 130408

QY 873 GATTCCTATGATGATGATTATGGAGATCCCTTCGAGAGGAACCTCTGCAGAAC 932  
Db 130409 GATTCCTATGATGATGATTATGGAGATCCCTTCGAGAGGAACCTCTGCAGAAC 130468

QY 933 CTTCAATACATTATGAATTTCCAG 957  
Db 130469 CTTCAATACATTATGAATTTCCAG 130493

RESULT 11  
US-10-238-709-3  
; Sequence 3, Application US/10238709  
; Patent No. 6680188  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEROP  
; FILE REFERENCE: CLO01164DIV  
; CURRENT APPLICATION NUMBER: US/10/238,709  
; CURRENT FILING DATE: 2002-09-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO. 3  
; LENGTH: 174493  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(174493)  
; OTHER INFORMATION: n = A,T,C or G

US-10-238-709-3

Query Match 3.3%; Score 205; DB 4; Length 174493;  
Best Local Similarity 100.0%; Pred. No. 5.7e-44;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 GGTGAATGCCAAACTCCCGATTGGGACCCAGATTACATGCTCTGAAGTCTGACTGT 812  
Db 130289 GGTGAATGCCAAACTCCCGATTGGGACCCAGATTACATGCTCTGAAGTCTGACTGT 130348

QY 813 GATGAACGGGGATGAAAGGACCTAGCGCTGGACTGTGACTGGTGGTGGGGCGT 872  
Db 130349 GATGAACGGGGATGAAAGGACCTAGCGCTGGACTGTGACTGGTGGTGGGGCGT 130408

QY 873 GATTCCTATGATGATGATTATGGAGATCCCTTCGAGAGGAACCTCTGCAGAAC 932  
Db 130409 GATTCCTATGATGATGATTATGGAGATCCCTTCGAGAGGAACCTCTGCAGAAC 130468

QY 933 CTTCAATACATTATGAATTTCCAG 957  
Db 130469 CTTCAATACATTATGAATTTCCAG 130493

RESULT 12  
US-08-630-822A-61  
; Sequence 61, Application US/08630822A  
; Patent No. 5840695  
; GENERAL INFORMATION:  
; APPLICANT: FRANK, GLENN R.

APPLICANT: HUNTER, SHIRLEY WU  
APPLICANT: WALLENFELS, LYNDIA  
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS  
TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,822A  
FILING DATE: 11-APR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CONNELL, GARY J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-17-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2706 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 5...2706  
US-08-630-822A-61

Query Match 3.1%; Score 189.2; DB 2; Length 2706;  
Best Local Similarity 54.1%; Pred. No. 7.9e-41;  
Matches 489; Conservative 0; Mismatches 388; Indels 27; Gaps 4;

QY 281 CAAAGGACTTCGAAAGTCAGAAAGTCTTGTAGTGTGTGTCTCACTTTTGTCTGAAAGTGCAGGTGG 340  
Db 78 CAGATGATTTAATTAAATTAAGTATTTGTCGAGGAGCATTTGGTGAAGTACAGTTAG 137

QY 341 TAAGAGAGAAAGCAACCGGGACATCTATGCTATGAAAGTATGAAAGAGAGAGCTTTAT 400  
Db 138 TCGACACAAATCAACTGCACAGTTTTTGTATGAAACGCCTATCAAAATTTGAAATGA 197

QY 401 TGGCCAGGAGCAGGTTTCATTTTTTTCGAGGAGAGCGGAACATATTTCTCGAAGCACA 460  
Db 198 TTAAGAGACAGACTCTGCATTTTTTTTGGGAGAGACGTCATATAATGGCCTCATCAAAAT 257

QY 461 GCCCGTGGATCCCCAATTACAGTATGCTTTTCAGGACAAAATCACTTTATCTGATGG 520  
Db 258 CAGAATGGATTGTACAATTACATTTTGTCTTCAAGATCAAAAATATCTTTATATGTGCA 317

QY 521 AGGAATATCAGCTGGAGGGGACTTGTGTCACATTTTGAATAGATATGAGGACCAAGTTAG 580  
Db 318 TGGATTATATCCGGGGGGTGTGCTGTGAGTCTTATG-----TCGATTATGAATTC 371

QY 581 ATGAAACCTGATACAGTTTTTACCTAGCTGAGCTGATTTTGGCTGTGTTTCACAGCGTTTCATC 640  
Db 372 CAGAAAATGGCAATGTTCTATACAATGGAAGTGGTGTAGCACTTTGATACAAATTCAT 431

QY 641 TGATGGGATAGCTGATCGAGACATCAAGCCTGAGACATTTCTGTTGACCGCAGGAC 700  
Db 432 CCATGGGATTTGTACATCGTGATGTTAAACCTGATAATATGCTTTCTAGACAAATATGGTC 491

QY 701 ACATCAAGCTGGTGAATTTTGGATCTGCGCGGAAATGAATTCACAAAGATGTTGAATG 760

Db 492 ATTTAAAGTTAGCTGACCTTTGGAAACCTGTATGAAATGGATACAGATGGTTTGGTAGCTT 551  
Qy 761 CCAAACTCCCGATTTGGGACCCAGATACATGGCTCTGAAAGTGGCTGACTGTGATGAACG 820  
Db 552 CTAATAATGCTGTTGGAAACGCTGATACATTTCTCCCGAAGT-----TTTGCAGTCCC 605  
Qy 821 GGGATGGAAGGACCACTAGCGCTGGAAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 880  
Db 606 AAGGTGGTGAAGGAGTTTACGGTCTGTAATGCGATTTGCTGCTGCTGCTGCTGCTGCTGCT 665  
Qy 881 ATGAGATGATTTATGGGAGATCCCTTTCCGAGAGGAACTCTGCGAGAACCTTCAATA 940  
Db 566 ATGAAATGTTATTTGGAGAACACCTTTTATGACAGACATTTGTTGGTGGAACTTACAGTA 725  
Qy 941 ACATTAATGAAATTTCCAGCGGTTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACT 1000  
Db 726 AAATTAATGATACAGAAACCTCAATTAACCTTTCTCCAGAAAGTGGAATAAGCCAATATG 785  
Qy 1001 TTCTTGATCTGATCAAGCTTTGTTGGCGCCAGAAAG-----AGACTGAAGT 1051  
Db 786 CCGCATCTTTGATACAGGATTTTAAACAGACATCCATTTTATGATGATCAATGACTTTTGA 845  
Qy 1052 TTGAAGGCTTTGCTGCTCCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1105  
Db 846 TGAAGAAATTAAGAGACATCCATTTTATGATGATCAATGACTTTTGAACAATTTAA 905  
Qy 1106 GTAACCT 1165  
Db 906 GAGACTCTGCCACCTGTAGTCCAGAGCTGAGTGTGATGATGATGATGATGATGATGATGAT 965  
Qy 1166 ATGA 1169  
Db 966 ATGA 969

RESULT 13

US-09-005-069-61  
; Sequence 61, Application US/09005069  
; Patent No. 5932470  
; GENERAL INFORMATION:  
; APPLICANT: FRANK, GLENN R.  
; APPLICANT: HUNTER, SHIRLEY WU  
; APPLICANT: WALLENFELS, LYNDIA  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS  
; TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross P.C.  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/005,069  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/630,822  
; FILING DATE: 11-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CONNELL, GARY J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-17-C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223

; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2706 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; NAME/KEY: CDS  
; LOCATION: 5..2706  
; US-09-005-069-61  
  
Query Match 3.1%; Score 189.2; DB 2; Length 2706;  
Best Local Similarity 54.1%; Pred. No. 7.9e-41;  
Matches 489; Conservative 0; Mismatches 388; Indels 27; Gaps 4;  
  
Qy 281 CAAAGGACTTCGAAGTCAGAAAGTCTTGTAGGTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340  
Db 78 CAGATGATTTTAAATTTTAAATAAGTTATTGTCAGGAGCAATTTGGTGAAGTACAGTTAG 137  
Qy 341 TAAGAGAGAAAGCAACCGGGGACATCTATGCTATGAAAGTCAATGAAGAAGAGGCTTTAT 400  
Db 138 TCGCAGACAAATCAACTGCAACAGTTTTTGTATGAAACGCTATCAAAATTTGAAATGA 197  
Qy 401 TGGCCCGAGGACGAGTTTTCATTTTTTGGAGAGAGCGGAACATATTTATCTCGAAGCAAA 460  
Db 198 TTAAGAGACCAAGCTCTGCAATTTTTTGGGAAGAACGTCATATAAATGGCTCATGCAAAAT 257  
Qy 461 CCCCCTGGATCCCCCAATTTACAGTATGCTTTCAGGACAAAAATCACCTTTTATCTGATGG 520  
Db 258 CAGAAATGATTTGTAATTTACATTTTCTTTCAGATCAAAATATCTTTATATGATGCA 317  
Qy 521 AGGAATATCAGCTGGAGGAGCTTGTGCTGCTCACTTTTGAATAGATAGAGGACAGTTAG 580  
Db 318 TGGATTATATGCGGGGGGTGACTTGTGTGATCTTATG-----TCGGATTATGAAATTC 371  
Qy 581 ATGAAACCTTGATACAGTTTTTACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCATC 640  
Db 372 CAGAAATGAGCAATGTTCTATACATGGAAGTGTGTAGCACTTGATCAATTCATC 431  
Qy 641 TGATGGGATAGCTGATCGAGAGACATCAAGCTCGAGAACATTCCTGTTGACCGCACAGGAC 700  
Db 432 CCATGGGATTTGTACATCGTGTGTTTAAACCTGATAATATGCTTCTAGACAAATATGGTC 491  
Qy 701 ACATCAAGCTGTGGATTTTGGATCTGCGCGGAAATGATTCAAACAAGATGCTGATG 760  
Db 492 ATTTAAAGTTAGCTGACTTTGGAACTGTATGAAATGGATACAGATGTTTGGTAGCTT 551  
Qy 761 CCAAACTCCCGATTTGGGACCCCGAGATTTACATGGCTCTCTGAAAGTCTGACTGTGATGAACG 820  
Db 552 CTAATAATGCTGTGGAAACGCTGATTACATTTCTCCGAAAT-----TTTGCAGTCCC 605  
Qy 821 GGGATGGAAGGACCACTAGCGGCTGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880  
Db 606 AAGGTGGTGAAGGAGTTTACGGTCTGTAATGCGATTTGTTGTTGCTGCTGCTGCTGCTGCT 665  
Qy 881 ATGAGATGATTTTATGGGAGATCCCTTTCCGAGAGGAACTCTGCGAGAACCTTCAATA 940  
Db 666 ATGAAATGTTATTTGGAGAAACACCTTTTATGACAGACATTTTGGTGGAACTTACAGTA 725  
Qy 941 ACATTATGAATTTCCAGCGGTTTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACT 1000  
Db 726 AAATTAATGATACAGAAATCAATTAACCTTTTCTCCAGAAAGTGAATAAGCCAATATG 785  
Qy 1001 TTCTTGATCTGATTCAAAGCTTCTTGTGGCGCCAGAAAG-----AGACTGAAGT 1051  
Db 786 CCGCATCTTTGATACAGGATTTTAAACAGACAGACACAGCGTTTGGCAGAAATGAAG 845  
Qy 1052 TTGAAGGCTTTGCTGCTGCT 1105  
Db 846 TGAAGAAATTAAGAGACATCCATTTTTCATAAATGATCAATGACTTTTGAACAATTTAA 905  
Qy 1106 GTAACT 1165

Db 906 GAGACTCTGCCACCTGTAGTCCAGAGCTGAGTGGTGATGATACAAAGAACTTTG 965  
QY 1166 ATGA 1169  
Db 966 ATGA 969

RESULT 14  
US-09-171-156A-20  
; Sequence 20, Application US/09171156A  
; Patent No. 6368846  
; GENERAL INFORMATION:  
; APPLICANT: Hunter, Shirley Wu  
; Weber, Eric R.  
; Sim, Gek-kee  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND  
; APPARATUS TO COLLECT SUCH PROTEINS  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHERIDAN ROSS P.C.  
; STREET: 1560 BROADWAY, SUITE 1200  
; CITY: DENVER  
; STATE: CO  
; COUNTRY: U.S.A.  
; ZIP: 80202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/171,156A  
; FILING DATE: 04-Mar-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 303/863-9700  
; TELEFAX: 303/863-0223  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2706 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 5..2706  
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-171-156A-20

Query Match 3.1%; Score 189.2; DB 4; Length 2706;  
Best Local Similarity 54.1%; Pred. No. 7.9e-41;  
Matches 489; Conservative 0; Mismatches 388; Indels 27; Gaps 4;

QY 281 CAAAGCACTCGAAGTCAGAAGTCTTGTAGTGTGGTCACTTGTCTGAAGTGCAGGTGG 340  
Db 78 CAGATGATTTTAATTTAATAAAGTTATTGGTCGAGGAGCAATTTGGTGAAGTACAGTTAG 137

QY 341 TAGAGAGAGAACCGGGACATCTATCTATGAAGTGTATGAAGAGAGAGGCTTTAT 400  
Db 138 TGGCAGACAAATCACTGCACAGTTTTTGTATGAAGCCCTATCAAAATTTGAATGA 197

QY 401 TGGCCCCAGGAGCAGGTTTCATTTTTTTCAGAGAGAGCGGAACATATTATCTCGAAGCACAA 460  
Db 198 TTAAGAGACCAGACTCTGCATTTTTTTTGGAGAGACGTCATATAATGGCTCATGCAAAAT 257

QY 461 GCCCTGGATCCCCCAATTACAGTATGCCCTTTCAGGACAAAAATCACCTTTATCTGATGG 520

Db 258 CAGAATGGATTGTACAATTACATTTTCTCTTCAAGATCAAAAATATCTTTATATGGTCA 317  
QY 521 AGGAATATCAGCTCGAGGGGACTTCTGTCTCACTTTTGAATAGATAGAGCACCACTTAG 580  
Db 318 TGGATTATATGCCGGGGGTGACTTGGTGAGTCTTAIG-----TCCGATTATGAAATTC 371

QY 581 ATGAAAACCTGATACAGTTTTTACCTAGCTAGCTGATTTGGCTGTTCACAGCGTTTCATC 640  
Db 372 CAGAAAAATGGCAATGTTCTATACAAATGGAAGTGGTGTAGCACTTGATACAAATCACT 431

QY 641 TGATGGGATACGTGCATCGAGACATCAAGCCTCAGAACATTTCTGTTGACCGCACAGGAC 700  
Db 432 CCATGGGATTTGACATCGTGATGTTAACTTGATATATGCTTCTAGACAAATATGGTC 491

QY 701 ACATCAAGCTGGTGGATTTTGGATCTGCCCGGAAAAATGAATTCAAAACAAGATGGTGAATG 760  
Db 492 ATTTAAAGTTAGCTGACTTTGGAAACCTGTATGAAAAATGGATACAGATGGTTTGGTACGTT 551

QY 761 CCAAACTCCGATTGGGACCCAGATTACATGCTCTCTGAAGTGTGACTGTGATGAACG 820  
Db 552 CTAATAATGCTGTTGGAAACCCCTGATTACATTTCTCCCGAAGT-----TTTGCAGTCCC 605

QY 821 GGGATGAAAAAGGCACCTACGGCCTCGACTGTGACTGTGCTGAGTGGTGGCGGTGATGCTCCT 880  
Db 606 AAGTGGTGAAGGAGTTTACGGTCTGATGCGATTGGTGGTCTGTGGGAATTTTITGT 665

QY 881 ATGAGATGATTTATGGAGATCCCTTCOCAGAGGAACTCTGCCAGAACCTTCCAGAACCTTCAATA 940  
Db 666 ATGAAATGTTATTGGGAAAAACACCTTTTATGCAGACAGTTTGGTTGGAACTTTACAGTA 725

QY 941 ACATTATGAATTTCCAGCGGTTTTTGAATTTCCAGATGACCCCAAGTGCAGCAGTCACT 1000  
Db 726 AATATATGATCACAGAACTCATTAACATTTTCTCCAGAAAGTGAATAAGCAATATG 785

QY 1001 TTCTGATCTGATCAAAAGCTTGTGGCGGCGCAGAAAGAG-----AGACTGAAGT 1051  
Db 786 CCGATCTTTGATACAGAGATTTTAAACAGACAGAACACACGCTTTAGGCAGAAATGAAG 845

QY 1052 TTGAAGGCTTTTGTGCTGCATCTTCTTCTCTAAATTTGACTGA-----ACAACATTC 1105  
Db 846 TGGAGAAATTAACAGACATCCATTTTTCATAATGATCAATGAGCTTTTGACAAATTTAA 905

QY 1106 GTAACCTCTCTCCCCCTTCTGTTCCACCTCAAGTCTGACGATGACACCTCCAAATTTTG 1165  
Db 906 GAGACTCTGCCCCACCTGTAGTCCAGAGCTGAGTGGTGTGATGATATACAAAGGAACCTTG 965

QY 1166 ATGA 1169  
Db 966 ATGA 969

RESULT 15  
US-09-004-730A-20  
; Sequence 20, Application US/09004730A  
; Patent No. 6485968  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eric  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Sim, Gek-kee  
; APPLICANT: Frank, Glenn  
; APPLICANT: Wallenfeils, Lynda  
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO COLLECT SUCH  
; FILE OF INVENTION: 2618-17-CS-PUS-1  
; CURRENT APPLICATION NUMBER: US/09/004,730A  
; CURRENT FILING DATE: 1998-01-08  
; PRIOR APPLICATION NUMBER: PCT/97US/18669  
; PRIOR FILING DATE: 1997-10-15  
; NUMBER OF SEQ ID NOS: 150  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 2706  
; TYPE: DNA  
; ORGANISM: Ctenocephalides felis

; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (5)..(2704)  
US-09-004-730A-20

Search completed: July 3, 2004, 19:34:05  
Job time : 286 secs

Query Match 3.1%; Score 189.2; DB 4; Length 2706;  
Best Local Similarity 54.1%; Pred.No. 7.9e-41;  
Matches 489; Conservative 0; Mismatches 388; Indels 27; Gaps 4;

QY 281 CAAAGGACTTCGAAGTCAGAAAGTCTTGTAGTCTGTGCTCACTTTGCTGAAGTGCAGGTGG 340  
DB 78 CAGATGATTTTAAATTAATTAAGATTATTGTCGAGGAGCAATTTGGTGAAGTACAGTTAG 137

QY 341 TAAGAGAGAAGCAACCGGGGACATCTATCTATGAAGTGAAGAGAGAGAGGCTTTTAT 400  
DB 138 TGGGACACAAATCAACTGCAAGTCTTGTCTATGAAGCGCTATCAAAATTTGAAATGA 197

QY 401 TGGCCAGGAGCAGGTTTCATTTTGGAGAGAGCGGAACATATTATCTCGAAGCACAA 460  
DB 198 TTAAGAGACCAGACTCTGCAATTTTGGGAAGAACTCATATATGCGCTCATGCAAAAT 257

QY 461 GCCCGTGGATCCCCCAATACAGTATGCCCTTTCAGGACAAAATCACCTTTATCTGATGG 520  
DB 258 CAGAATGGATTGACAAATACATTTGCTTTCAAGATCAAAAATATCTTTATATGCTCA 317

QY 521 AGGAATATACAGCTGGAGGGGACTTGTCTCACTTTTGAATAGATATGAGGACCAGTTAG 580  
DB 318 TGGATTATATGCCGGGGGTGACTTGGTGAAGCTTTATG-----TCCGATTATGAAATTC 371

QY 581 ATGAAACCTGTATACAGTTTACCTAGCTGAGCTGATTTTGGCTGTTCACAGCTTCATC 640  
DB 372 CAGAAAAATGGCAATGTTCTATACAAATGGAAGTGTGCTAGCACTTGATACAAATCACT 431

QY 641 TGATGGATACGTGCAATCGAGACATCAAGCTCGAGAACATTTCTGTTGACCGCACAGGAC 700  
DB 432 CCATGGGATTGACATCGTGATGTTTAAACCTGATAATATGCTTCTAGACAAATATGCTC 491

QY 701 ACATCAAGCTGTGTGGATTTTGGATCTGCCGGAATAATGAAATCAAAACAGATGTTGAATG 760  
DB 492 ATTAAAGTTAGCTGACTTTGGAACCTGTATGAAATGGATACAGATGGTTTGGTACGTT 551

QY 761 CCAAATCCGGATTGGGACCCGAGATTAATGCGCTCTCGAAGTGTGACTGTGATGAACG 820  
DB 552 CTAATAATGCTGTGGAAACGCTGATTACATTTCTCCCGAAGT-----TTTGCAGTCCC 605

QY 821 GGGATGGAAGGACCTACGGCTGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880  
DB 606 AAGTGGTGAAGAGTTTACGGTCTGGAATGCGATTGGTGGTCTGTTGGGAATTTTITGT 665

QY 881 ATGAGATGATTTATGGAGATCCCTCTCGCAGAGGGAACCTCTGCCAGAACCTTCAATA 940  
DB 666 ATCAATGTTATTTGGAGAAACACCTTTTATGACAGACAGTTTGGTGGAACTTACAGTA 725

QY 941 ACATATGATTTCCAGCGGTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGT 1000  
DB 726 AAATATGATCACAGAACTCAATTAATTTTCTCCAGAGTGAAGTGAAGTGAAGTGAAG 785

QY 1001 TTCTTGATCTGATCAAAAGCTTTGTTGTGGCGCCAGAAAGAG-----AGACTGAAGT 1051  
DB 786 CCCGATCTTTGATACAAAGATTTTAAACAGACAGAACACAGCGTTTAGGCAGAAATGAAG 845

QY 1052 TTGAAGGCTTTTGTGCGCACTTCTTCTCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1105  
DB 846 TGGAAAGAAATTAACAGACATCCATTTTCAATTAATGATCAATGGACTTTTGACAAATTTAA 905

QY 1106 GTAACCTCTCTCCCTCCCTCTGTTCCACCTCAAGCTCTGACGATGACACCTCCCAATTTTG 1165  
DB 906 GAGACTCTGCCCCACCTCTAGTGGCCAGAGCTGAGTGGTGAATGATGATACAGGAACCTTG 965

QY 1166 ATGA 1169  
DB 966 ATGA 969

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2004, 10:49:43 ; Search time 1455 Seconds  
(without alignments)  
17982.579 Million cell updates/sec

Title: US-10-017-216-3

Perfect score: 6159  
Sequence: 1 atgttgaagtccaatatgg.....ttctgagaacagattattgc 6159

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002s:\*
- 7: Geneseqn2003as:\*
- 8: Geneseqn2003bs:\*
- 9: Geneseqn2003cs:\*
- 10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6159	100.0	6574	6	AAD39191 Human MDP
2	5668	92.0	8603	8	AAL55215 Human CRI
3	5666	92.0	6156	8	AAL55217 Human CRI
4	5666	92.0	6298	6	AAD38864 Human CRI
5	5666	92.0	6165	8	AAL55214 Human CRI
6	5661	91.9	6165	6	ABQ78870 Human kin
7	5650	91.4	6159	6	AAS06701 Polynucle
8	5631	91.4	6189	6	ABS63436 RHO/RAC-1
9	5631	91.4	6189	6	ADA05641 Human NOV
10	5629	91.4	6201	6	ABS63435 Human cdn
11	5629	91.4	6201	7	ADA05653 Human NOV
12	5358	87.0	5877	6	ABQ78871 Human kin
13	3475	56.4	6609	3	AAC77568 Human ORF
14	2475	41.8	2896	5	ABV30132 Human pro
15	2564	41.6	3131	4	ABA08361 Human RHO
16	2436	39.6	5261	8	AAL55216 Human CRI
17	2426	39.4	5251	9	ADD89966 Human can
18	2414	39.2	2542	7	ADA05647 Human NOV
19	2317	37.6	2437	7	ADA05645 Human NOV
20	2126	34.5	2693	9	ADE09823 Novel DNA
21	1841	29.9	1870	7	ADA05643 Human NOV
22	1524	24.7	1870	7	ADA05649 Human NOV
23	1467	23.8	1915	7	ADA05651 Human NOV

24	1395	22.6	1485	7	ABZ68725	Abz68725 Nucleotid
25	1395	22.6	1755	7	ABZ68726	Abz68726 Nucleotid
26	1393	22.6	2066	6	AAD26454	Aad26454 Human kin
27	1391	22.6	1515	7	ACA61394	Aca61394 cDNA enco
28	1391	22.6	1515	9	AAD59938	Aad59938 Human kin
29	1082	17.6	2380	7	ABZ68776	Abz68776 Nucleotid
30	952	15.5	2162	7	ABX71191	Abx71191 Novel hum
31	930	15.1	1048	5	AAS79753	Aas79753 DNA encod
32	701	11.4	995	4	ABA08479	Aba08479 Human cit
33	648	10.5	1058	4	AAP22502	Aaf22502 Human bre
34	592	9.6	817	4	AAP22603	Aaf22603 Human bre
35	443	7.2	446	5	ABV15823	Abv15823 Human pro
36	436	7.1	485	5	ABV45624	Abv45624 Human pro
37	433	7.0	580	4	AAL192420	Aai192420 Human pol
38	350	5.7	354	2	AAV87531	Aav87531 EST clone
39	337	5.5	396	4	AAH99003	Aah99003 Marine ES
40	258	4.2	258	7	ACA55915	Aca55915 Mouse sig
41	250	4.1	254	2	AAV89320	Aav89320 EST clone
42	244	4.0	3835	7	ABT33346	Abt33346 NOX DNA
43	244	4.0	3985	7	ABT33347	Abt33347 NOX DNA
44	243	4.0	1530	6	AAD38845	Aad38845 Human kin
45	243	4.0	5373	6	AAD30567	Aad30567 Human kin

## ALIGNMENTS

RESULT 1

AAD39191  
ID AAD39191 standard; cDNA; 6574 BP.

XX AC AAD39191;

XX DT 04-OCT-2002 (first entry)

XX DE Human MDPK cDNA.

XX KW Human; myotonic dystrophy type protein kinase; MDPK; 13245 protein;  
tumorigenesis; tumour growth; tumour metastasis; viral infection;  
skeletal muscle disorder; muscular dystrophy; myotonic dystrophy;  
immune disorder; neoplastic disorder; gene therapy; gens; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT 5'UTR 1..18

FT CDS /\*tag= a

FT CDS 19..6180

FT misc\_feature /\*tag= b

FT misc\_feature 19..6177

FT misc\_feature /\*tag= c

FT 3'UTR /note= "This region is specifically referred in claim 1  
as SEQ ID NO:3"

FT 3'UTR 6181..6574

FT 3'UTR /\*tag= d

XX WO200234896-A2.

XX PN Aas06701 Polynucle

XX PD Abs63436 RHO/RAC-1

XX PD Ada05641 Human NOV

XX PD Abs63435 Human cdn

XX PF Ada05653 Human NOV

XX PF Abq78871 Human kin

XX PR Aac77568 Human ORF

XX PR Abv30132 Human pro

XX XX ABA08361 Human RHO

XX PA AAL55216 Human CRI

XX PI ADD89966 Human can

XX PI Ada05647 Human NOV

XX XX ADE09823 Novel DNA

XX DR WPI; 2002-479720/51.

XX DR P-FSDB; AAE24079.

XX PT Human myotonic dystrophy type protein kinase polypeptide and

PT polynucleotide useful for prognosticating, diagnosing, preventing or  
PT inhibiting tumorigenesis, tumor growth, tumor metastasis and viral  
XX infection.

XX Claim 1; Fig 1; 148pp; English.

XX CC The invention relates to human myotonic dystrophy type protein kinase  
CC (MDPK) polypeptides designated as 13245 and nucleic acid molecules  
CC encoding such polypeptides. 13245 molecules are used to develop  
CC diagnostic and therapeutic agents for prognosticating, diagnosing,  
CC preventing, inhibiting, alleviating or curing MDPK-related disorders.  
CC Polypeptides of the invention are used to develop diagnostic and  
CC therapeutic agents for 13245-mediated or related disorders such as  
CC tumorigenesis, tumour growth, tumour metastasis, viral infection of a  
CC cell, skeletal muscle disorders (e.g. muscular and myotonic dystrophies),  
CC immune disorders and neoplastic disorders. The invention is also used in  
XX gene therapy. The present sequence is human MDPK cDNA

SQ Sequence 6574 BP; 1877 A; 1611 C; 1776 G; 1310 T; 0 U; 0 Other;

Query Match 100.0%; Score 6159; DB 6; Length 6574;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 6159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGGTCTCTGAACCCCAAT	60
DB	19	ATGTTGAAGTTCAAAATATGAGCGCGGAATCCCTTTGGATGCTGGTCTCTGAACCCCAAT	78
QY	61	GCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGAACACCCTTTATGATCTCAA	120
DB	79	GCAGCGCGGCTCCAGGCTGAATCTGTTCTCCAGGGAACACCCTTTATGATCTCAA	138
QY	121	CAGCAGATGTCCTCTTTCCCGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA	180
DB	139	CAGCAGATGTCCTCTTTCCCGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA	198
QY	181	GAATGAGTCAGCTGCTCTGATGAAGATTAGACAGTGAAGAACTTTCCGGAGTAT	240
DB	199	GAATGAGTCAGCTGCTCTGATGAAGATTAGACAGTGAAGAACTTTCCGGAGTAT	258
QY	241	TCGCACACCATAGCTCAGTTACAGGAGCTCCAGCCTTCGCAAGAGCACTTCCAGTCTAGA	300
DB	259	TCGCACACCATAGCTCAGTTACAGGAGCTCCAGCCTTCGCAAGAGCACTTCCAGTCTAGA	318
QY	301	AGTCTTGTAGTTGGTCACTTTGCTGAAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG	360
DB	319	AGTCTTGTAGTTGGTCACTTTGCTGAAGTGCAGGTGTTAAGAGAGAAAGCAACCGGG	378
QY	361	GACATCTATGCTATGAAGTATGAGAGAGAGGCTTTTATTCGCCCGAGGAGGTTTCA	420
DB	379	GACATCTATGCTATGAAGTATGAGAGAGAGGCTTTTATTCGCCCGAGGAGGTTTCA	438
QY	421	TTTTTTGAGAAAGAGCGGAACATATTATCTGAAGCACAAGCCCGTGGATCCCCCAATTA	480
DB	439	TTTTTTGAGAAAGAGCGGAACATATTATCTGAAGCACAAGCCCGTGGATCCCCCAATTA	498
QY	481	CAGTATGCCTTTCAGACAAAATACCTTTATCTGATGGAGGAATATCAGCCTGGAGG	540
DB	499	CAGTATGCCTTTCAGACAAAATACCTTTATCTGATGGAGGAATATCAGCCTGGAGG	558
QY	541	GACTTCTGCTCACTTTTGAATAGATATGAGGACCACTTATAGTGAATACTGTATCAGTTT	600
DB	559	GACTTCTGCTCACTTTTGAATAGATATGAGGACCACTTATAGTGAATACTGTATCAGTTT	618
QY	601	TACCTAGCTAGCTGATTTTGGCTGTTTCAACGGTTTCACTGATGGGATACGTGATCGA	660
DB	619	TACCTAGCTAGCTGATTTTGGCTGTTTCAACGGTTTCACTGATGGGATACGTGATCGA	678
QY	661	GACATCAAGCTTGAGAACATTTCTCGTTGACCGCACAGGACATCAAGCTGTGATTTT	720
DB	679	GACATCAAGCTTGAGAACATTTCTCGTTGACCGCACAGGACATCAAGCTGTGATTTT	738
QY	721	GGATCTGCCGGAATAATGAATCAAAACAGATGGTGAATGCGAAACTCCCGATTTGGGACC	780

DB	739	GGATCTGCCGGAATAATGAATCAAAACAGATGGTGAATGCGAAACTCCCGATTTGGGACC	798
QY	781	CCAGATTACATGGCTCTCTGAAGTCTGACTGTGATGAACCGGGATGGAAGGACCTAC	840
DB	799	CCAGATTACATGGCTCTCTGAAGTCTGACTGTGATGAACCGGGATGGAAGGACCTAC	858
QY	841	GGCTTGACTGTGACTGGTGTGAGTGGGGGTGATTCCTATGAGATGATTTATGGGAGA	900
DB	859	GGCTTGACTGTGACTGGTGTGAGTGGGGGTGATTCCTATGAGATGATTTATGGGAGA	918
QY	901	TCCCTCTTCGAGAGGAACTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG	960
DB	919	TCCCTCTTCGAGAGGAACTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG	978
QY	961	TTTTTGAATAATTCAGATGACCCCAAGAGTGAAGTGAATTTCTTGATCTGATTTCAAAGC	1020
DB	979	TTTTTGAATAATTCAGATGACCCCAAGAGTGAAGTGAATTTCTTGATCTGATTTCAAAGC	1038
QY	1021	TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC	1080
DB	1039	TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTCTTC	1098
QY	1081	TCTAAAATTGACTGGAAACAATTCGTAATCTCTCTCCCTCCCTTCGTTCCACCTCAAG	1140
DB	1099	TCTAAAATTGACTGGAAACAATTCGTAATCTCTCTCCCTCCCTTCGTTCCACCTCAAG	1158
QY	1141	TCTGACCATGACACTCCAAATTTTGATGAACCGAGAGAAATTCGTGGGTTTCATCTCT	1200
DB	1159	TCTGACCATGACACTCCAAATTTTGATGAACCGAGAGAAATTCGTGGGTTTCATCTCT	1218
QY	1201	CCGTGCCAGTGAAGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG	1260
DB	1219	CCGTGCCAGTGAAGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTCG	1278
QY	1261	TACAGCAAGGACTGGGGATTTTGGTAGATCTGAGTCTGTTGTGCGGTCTGGACTCC	1320
DB	1279	TACAGCAAGGACTGGGGATTTTGGTAGATCTGAGTCTGTTGTGCGGTCTGGACTCC	1338
QY	1321	CCTGCCAAGTACTGCTCCATGGAAGAAATCTCTCATCAAAAGCAAGAGCTACAAGAC	1380
DB	1339	CCTGCCAAGTACTGCTCCATGGAAGAAATCTCTCATCAAAAGCAAGAGCTACAAGAC	1398
QY	1381	TCTCAGCAAGTGTCAAGATGAGAGAGAAATGACCGGTTTACATCGGAGGTGTCA	1440
DB	1399	TCTCAGCAAGTGTCAAGATGAGAGAGAAATGACCGGTTTACATCGGAGGTGTCA	1458
QY	1441	GAGTGGAGGCTGTGCTTAGTCAGAGGAGTGAAGCTGAAGGCTCTGAGACTCAGAGA	1500
DB	1459	GAGTGGAGGCTGTGCTTAGTCAGAGGAGTGAAGCTGAAGGCTCTGAGACTCAGAGA	1518
QY	1501	TCCCTCTCGAGCAGGACCTTGTCTAATCATCAAGATGCAAGTGTAAAGCGAAGT	1560
DB	1519	TCCCTCTCGAGCAGGACCTTGTCTAATCATCAAGATGCAAGTGTAAAGCGAAGT	1578
QY	1561	TTGGAGCAGCAGGATGGAGTGTCCAGGAGGATGCAAAAGCACTGAGCTTCTCCAT	1620
DB	1579	TTGGAGCAGCAGGATGGAGTGTCCAGGAGGATGCAAAAGCACTGAGCTTCTCCAT	1638
QY	1621	GATATCAGAGCAGAGCCCGAAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGCTCAA	1680
DB	1639	GATATCAGAGCAGAGCCCGAAAGCTCCAAAGAAATCAAGAGCAGGAGTACCAAGCTCAA	1698
QY	1681	GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1740
DB	1699	GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1758
QY	1741	CGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTCTCAAGAAATCAAG	1800
DB	1759	CGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTCTCAAGAAATCAAG	1818
QY	1801	CGAAAGCGCAGAAATGTGACGATAAATCTGTTGAAGCTTAAGATCAAGGAGGCTGAA	1860

Db 1819 CGGAAGCGACAGAATGTGACGATAAACTGTTGAAGCTTAAGGATCAAGGGAAGCCTGAA 1878  
QY 1861 GTGGAGAAATATGCGAAATCGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
Db 1879 GTGGAGAAATATGCGAAATCGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1938  
QY 1921 CTCGAAGAAACTCGAGAGGCTGCAAGAGCGTCAAGAGCGAGCGCGAGAGGAGCTGGAAAGCTG 1980  
Db 1939 CTCGAAGAAACTCGAGAGGCTGCAAGAGCGTCAAGAGCGAGCGCGAGAGGAGCTGGAAAGCTG 1998  
QY 1981 CAGAACCGAGAGGATCTCTTCTCAAGGCATCAAGAAAGAGCTGGTGAAGCTGAGGAAGCG 2040  
Db 1999 CAGAACCGAGAGGATCTCTTCTCAAGGCATCAAGAAAGAGCTGGTGAAGCTGAGGAAGCG 2058  
QY 2041 CCGCATCTCTCGAGAAACAAGTAAAGAGACTAGAGACCATGAGCGCTAGAGAAACAGA 2100  
Db 2059 CCGCATCTCTCGAGAAACAAGTAAAGAGACTAGAGACCATGAGCGCTAGAGAAACAGA 2118  
QY 2101 CTGAAGGATGATCCAGACAAATCCCAACAGATCCAGCAGATCGCTGATAAAATTCGT 2160  
Db 2119 CTGAAGGATGATCCAGACAAATCCCAACAGATCCAGCAGATCGCTGATAAAATTCGT 2178  
QY 2161 GAGCTCGAAGAAACATCGGAGGCCCAAGTCTCAGGCCAGACCTAGAGTGCACCTG 2220  
Db 2179 GAGCTCGAAGAAACATCGGAGGCCCAAGTCTCAGGCCAGACCTAGAGTGCACCTG 2238  
QY 2221 AAACGAAAGAGAGCACTATGAGGAAGAAGATTAAGTGTGGAACAATCAGATAAAGAAA 2280  
Db 2239 AAACGAAAGAGAGCACTATGAGGAAGAAGATTAAGTGTGGAACAATCAGATAAAGAAA 2298  
QY 2281 GACCTGGCTGCAAGAGAGACATCGGAGAACATGATGCAGAGACACGAGGAGAGGCCCAT 2340  
Db 2299 GACCTGGCTGCAAGAGAGACATCGGAGAACATGATGCAGAGACACGAGGAGAGGCCCAT 2358  
QY 2341 GAGAGGGCAAAATTTCTCAGGAAACAGAGGCGATGATCAATGCTATGATTTCCAGATC 2400  
Db 2359 GAGAGGGCAAAATTTCTCAGGAAACAGAGGCGATGATCAATGCTATGATTTCCAGATC 2418  
QY 2401 AGATCCCTGGACAGAGGATTTGTGAACTGTCTGAAGCCAAATAAATTCGACCAATAGC 2460  
Db 2419 AGATCCCTGGACAGAGGATTTGTGAACTGTCTGAAGCCAAATAAATTCGACCAATAGC 2478  
QY 2461 AGTCTTTTTTACCAGAGAACATGAAGGCCCAAGAGAGAGATGATTTCTGAATCTCAGGCAA 2520  
Db 2479 AGTCTTTTTTACCAGAGAACATGAAGGCCCAAGAGAGAGATGATTTCTGAATCTCAGGCAA 2538  
QY 2521 CAGAAATTTTACTGTGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAATACTGGAG 2580  
Db 2539 CAGAAATTTTACTGTGAGACACAGGCTGGGAAGTTGGAGGCCAGAACCGAATACTGGAG 2598  
QY 2581 GAGCAGCTGGAGAAAGATCAGCCACCAAGACCACAGTGACAAGATCGGCTGTGGAACCTG 2640  
Db 2599 GAGCAGCTGGAGAAAGATCAGCCACCAAGACCACAGTGACAAGATCGGCTGTGGAACCTG 2658  
QY 2641 GAGACAAGATTGCGGAGGTCAGTCTAGAGCAAGAGGAGCAGAAATCGGAGCTCAAGCGC 2700  
Db 2659 GAGACAAGATTGCGGAGGTCAGTCTAGAGCAAGAGGAGCAGAAATCGGAGCTCAAGCGC 2718  
QY 2701 CAGCTCACAGAGCTACAGCTCTCCCTGCAAGGACCGGAGTCAAGTTGACGCCCTGCAG 2760  
Db 2719 CAGCTCACAGAGCTACAGCTCTCCCTGCAAGGACCGGAGTCAAGTTGACGCCCTGCAG 2778  
QY 2761 GCTGCACGGCGGCCCTGGAGAGCCAGCTTCGCGAGCGGAAGACAGAGCTGGAAGAGACC 2820  
Db 2779 GCTGCACGGCGGCCCTGGAGAGCCAGCTTCGCGAGCGGAAGACAGAGCTGGAAGAGACC 2838  
QY 2821 ACAGCAAGAGCTGAAAGAGGAGATCCAGGCACTCAGCGCAATAGAGATGAATTCAGCGC 2880  
Db 2839 ACAGCAAGAGCTGAAAGAGGAGATCCAGGCACTCAGCGCAATAGAGATGAATTCAGCGC 2898  
QY 2881 AAATTTGATCCTCTCGTAAACAGCTGTACTGTAAATCACAGACTCGGAGGAGCAGCTAAAC 2940  
Db 2899 AAATTTGATCCTCTCTGTAAACAGCTGTACTGTAAATCACAGACTCGGAGGAGCAGCTAAAC 2958

QY 2941 CAGCTGACCGAGGACAAACGCTGAACCTCAACAAACAAACCTTCTACTTGTCCAAACAATC 3000  
Db 2959 CAGCTGCCGAGGACAAACGCTGAACCTCAACAAACAAACCTTCTACTTGTCCAAACAATC 3018  
QY 3001 GATGAGGCTTCTCGCGCCACGACGAGATGTGTAACCTGCGAAGTGAAGTGGACCATCTC 3060  
Db 3019 GATGAGGCTTCTCGCGCCACGACGAGATGTGTAACCTGCGAAGTGAAGTGGACCATCTC 3078  
QY 3061 CGCGGAGATCACGGAAACGAGAGATGACGCTTTACAGCCAGAAACGATGGAGGCT 3120  
Db 3079 CGCGGAGATCACGGAAACGAGAGATGACGCTTTACAGCCAGAAACGATGGAGGCT 3138  
QY 3121 CTGAAGACACGCTGCACCATCTGGAGGAACAGGTGATGATTTGGAGGCCCTTAAACGAT 3180  
Db 3139 CTGAAGACACGCTGCACCATCTGGAGGAACAGGTGATGATTTGGAGGCCCTTAAACGAT 3198  
QY 3181 GAGCTGCTAGAAAAGAGCGCAGTGGAGGCCCTGGAGGAGCCTCTCTGGGTGATGAGAAA 3240  
Db 3199 GAGCTGCTAGAAAAGAGCGCAGTGGAGGCCCTGGAGGAGCCTCTCTGGGTGATGAGAAA 3258  
QY 3241 TCCAGATTTGAGTGTTCGGGTTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGC 3300  
Db 3259 TCCAGATTTGAGTGTTCGGGTTTCGAGAGCTGCAGAGAAATGCTGGACACCGAGAAACAGAGC 3318  
QY 3301 AGGCGGAGAGCCGATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAG 3360  
Db 3319 AGGCGGAGAGCCGATCAGCGGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAG 3378  
QY 3361 GAGCACAAAGCTGAGATTTCTCGCTCTCCAGCAGGCTCTCAAAGAGCAGAGCTGAAAGGCC 3420  
Db 3379 GAGCACAAAGCTGAGATTTCTCGCTCTCCAGCAGGCTCTCAAAGAGCAGAGCTGAAAGGCC 3438  
QY 3421 GAGAGCCTCTCTGCAAGCTCAATGACCTGGAGGAAGAGCATGCTATGCTTGAATGAAT 3480  
Db 3439 GAGAGCCTCTCTGCAAGCTCAATGACCTGGAGGAAGAGCATGCTATGCTTGAATGAAT 3498  
QY 3481 GCCCGAAGCTTACAGCAGAAAGCTGGAGACTCAACGAGCTCAAAACAGAGGCTTCTCGAA 3540  
Db 3499 GCCCGAAGCTTACAGCAGAAAGCTGGAGACTCAACGAGCTCAAAACAGAGGCTTCTCGAA 3558  
QY 3541 GAGCAAGCCAAATTAAGCAGCAGATGGAGCTGCAGAAAAATCAATTTCCGCTGACT 3600  
Db 3559 GAGCAAGCCAAATTAAGCAGCAGATGGAGCTGCAGAAAAATCAATTTCCGCTGACT 3618  
QY 3601 CAAGGACTGCAAGAAAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3660  
Db 3619 CAAGGACTGCAAGAAAGCTTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAAGCTG 3678  
QY 3661 GAGTATCAGCTGGAAAAATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAAATGGAAGGC 3720  
Db 3679 GAGTATCAGCTGGAAAAATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAAATGGAAGGC 3738  
QY 3721 ACTATTTCTCAACAAACCAACTCATTGATTTTCTGCAAGCCAAATGGACCAACCTGCT 3780  
Db 3739 ACTATTTCTCAACAAACCAACTCATTGATTTTCTGCAAGCCAAATGGACCAACCTGCT 3798  
QY 3781 AAAAAGAAAAAGGTTTTATTAGTCGACGGAAGAGGACCTGCTTTACCCACACAGGTT 3840  
Db 3799 AAAAAGAAAAAGGTTTTATTAGTCGACGGAAGAGGACCTGCTTTACCCACACAGGTT 3858  
QY 3841 CCTCTGAGTACAATGAGCTGAAGCTGGCCCTGGAGGAAGAGAAAGCTCGCTGTGCAGAG 3900  
Db 3859 CCTCTGAGTACAATGAGCTGAAGCTGGCCCTGGAGGAAGAGAAAGCTCGCTGTGCAGAG 3918  
QY 3901 CTAGAGGAAGCCCTTCAGAGACCCGATCGAGCTCCGGTCCGGCCCGGAGGAAGCTGCC 3960  
Db 3919 CTAGAGGAAGCCCTTCAGAGACCCGATCGAGCTCCGGTCCGGCCCGGAGGAAGCTGCC 3978  
QY 3961 CACCGCAAGCAACGAGACCCACACCCATCCAGCCAGCCAGCCAGCGAGGAGCAGATC 4020  
Db 3979 CACCGCAAGCAACGAGACCCACACCCATCCAGCCAGCCAGCCAGCGGAGGAGCAGATC 4038

QY 4021 GCCATGTCGCCATCGTGGCTCGCCAGAGCACCAGCCAGTGCCTGAGCCTGCTGGCC 4080  
Db 4039 GCCATGTCGCCATCGTGGCTCGCCAGAGCACCAGCCAGTGCCTGAGCCTGCTGGCC 4098  
QY 4081 CCGCCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTTTAAG 4140  
Db 4099 CCGCCATCCAGCCGAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTTTAAG 4158  
QY 4141 GAAGCATGACACACATATTCCTCCAGATTCAAGCTAGGAGTGAACATGGAGCCACA 4200  
Db 4159 GAAGCATGACACACATATTCCTCCAGATTCAAGCTAGGAGTGAACATGGAGCCACA 4218  
QY 4201 AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGACGCCAGGCATCCAAATGCTTCGAA 4260  
Db 4219 AAGTGTGCTGTGTCTGGATACCGTGCACTTTGGACGCCAGGCATCCAAATGCTTCGAA 4278  
QY 4261 TGTCAAGTGTGTGTCAACCCAGTGTCTGCAAGTGTCTGCGAGCCACTTGGGCTTGGCT 4320  
Db 4279 TGTCAAGTGTGTGTCAACCCAGTGTCTGCAAGTGTCTGCGAGCCACTTGGGCTTGGCT 4338  
QY 4321 GCTGAATATGCCACACATTTACCGAGGCTTCTGCGGTGAACAAATGAACCTCCACAGT 4380  
Db 4339 GCTGAATATGCCACACATTTACCGAGGCTTCTGCGGTGAACAAATGAACCTCCACAGT 4398  
QY 4381 CTCAGACCAAGAGCCAGCAGCAGCTTGCACCTGGAAAGGTGGATGAAGTGGCCAGG 4440  
Db 4399 CTCAGACCAAGAGCCAGCAGCAGCTTGCACCTGGAAAGGTGGATGAAGTGGCCAGG 4458  
QY 4441 AATAACAAACGAGCAGCAAGGCTGGACAGGAAGTACATTTGCTCGAGGATCAAAA 4500  
Db 4459 AATAACAAACGAGCAGCAAGGCTGGACAGGAAGTACATTTGCTCGAGGATCAAAA 4518  
QY 4501 GTCTCATTTATGACAAATGAAGCAGAGAGCTGGAACAGCCCGGTGGAAGAAATTTGAG 4560  
Db 4519 GTCTCATTTATGACAAATGAAGCAGAGAGCTGGAACAGCCCGGTGGAAGAAATTTGAG 4578  
QY 4561 CTGTGCTTTCCACAGGGATGTATCTATTCATGTTGTCGCTGGTGGTCTCCGAACTCGCA 4620  
Db 4579 CTGTGCTTTCCACAGGGATGTATCTATTCATGTTGTCGCTGGTGGTCTCCGAACTCGCA 4638  
QY 4621 AATACAGCCAAAGCAGAAAGCAGAAAGCTGATGCTTAAACTGTGTTGAAACTCCCTGTG 4680  
Db 4639 AATACAGCCAAAGCAGAAAGCAGAAAGCTGATGCTTAAACTGTGTTGAAACTCCCTGTG 4698  
QY 4681 AACTGGAGGTGATGACCGTCTAGACATGAACCTGACCGCTGCTTCTGAGTACAGGTG 4740  
Db 4699 AACTGGAGGTGATGACCGTCTAGACATGAACCTGACCGCTGCTTCTGAGTACAGGTG 4758  
QY 4741 GTCTTGGTGGCACCAGGAAGGCTCTACGCCCTGAATGTCTTGAATAACTCCCTAACC 4800  
Db 4759 GTCTTGGTGGCACCAGGAAGGCTCTACGCCCTGAATGTCTTGAATAACTCCCTAACC 4818  
QY 4801 CATGTCCCAGGAATTTGGAGAGTCTTCCAAATTTATATTTATCAAGAACTGGAAGCTA 4860  
Db 4819 CATGTCCCAGGAATTTGGAGAGTCTTCCAAATTTATATTTATCAAGAACTGGAAGCTA 4878  
QY 4861 CTCATGATAGCAGGAAGAGCGGCACGTGCTTGTGGAGCTGAAGAAAGTGAACAG 4920  
Db 4879 CTCATGATAGCAGGAAGAGCGGCACGTGCTTGTGGAGCTGAAGAAAGTGAACAG 4938  
QY 4921 TCCTGGCCAGTCCACCTGCTGCCAGCCCGACATCTCACCAACATTTTGAAGCT 4980  
Db 4939 TCCTGGCCAGTCCACCTGCTGCCAGCCCGACATCTCACCAACATTTTGAAGCT 4998  
QY 4981 GTCAAGGCTGCACCTGTTTGGGACGCGAGATTTGAAACGGCTCTGATCTGTGCA 5040  
Db 4999 GTCAAGGCTGCACCTGTTTGGGACGCGAGATTTGAAACGGCTCTGATCTGTGCA 5058  
QY 5041 GCCATGCCAGCAAAAGTCTGATTTCTCCGCTACACGAAACCTCAGCAAAATTAAGTATC 5100  
Db 5059 GCCATGCCAGCAAAAGTCTGATTTCTCCGCTACACGAAACCTCAGCAAAATTAAGTATC 5118  
QY 5101 CGGAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCACTTACCAATTAAGTATC 5160

Db 5119 CGGAAGAGATAGAGACCTCAGAGCCCTGAGCTGTATCACTTACCAATTAAGTATC 5178  
QY 5161 CTCATTGGAAACCAATAAATTCTACGAAATGACATGAAGCAGTACAGCTCGAGGAATTC 5220  
Db 5179 CTCATTGGAAACCAATAAATTCTACGAAATGACATGAAGCAGTACAGCTCGAGGAATTC 5238  
QY 5221 CTGATTAAGAAATGACCAATTCCTTGGCCACTGCTGTGTTTGGCCCTCTTCCAAACAGCTTC 5280  
Db 5239 CTGATTAAGAAATGACCAATTCCTTGGCCACTGCTGTGTTTGGCCCTCTTCCAAACAGCTTC 5298  
QY 5281 CCTGCTCAATCGTGCAGGTGAACAGCGCAGGGCAGCAGAGAGTACTTGTCTGTGTTTC 5340  
Db 5299 CCTGCTCAATCGTGCAGGTGAACAGCGCAGGGCAGCAGAGAGTACTTGTCTGTGTTTC 5358  
QY 5341 CACGAATTTGAGTGTTCGTGGATTCTTACGGAAGAGTACGCGCAGAGAGTACTTGTCTGTGTTTC 5400  
Db 5359 CACGAATTTGAGTGTTCGTGGATTCTTACGGAAGAGTACGCGCAGAGAGTACTTGTCTGTGTTTC 5418  
QY 5401 TGGAGTGTCTTACCTTTGGCCCTTACAGAAACCTTATCTGTTGTGACCACTTC 5460  
Db 5419 TGGAGTGTCTTACCTTTGGCCCTTACAGAAACCTTATCTGTTGTGACCACTTC 5478  
QY 5461 AACTCACTCGAAGTAATGAGATCCAGGACGCTCTCAGAGGAGCCCTGCCCCAGCG 5520  
Db 5479 AACTCACTCGAAGTAATGAGATCCAGGACGCTCTCAGAGGAGCCCTGCCCCAGCG 5538  
QY 5521 TACTGACATCCCGAACCCCGCTACTGCGGCCCTGCCATTTCTCAGGAGGAGTATAC 5580  
Db 5539 TACTGACATCCCGAACCCCGCTACTGCGGCCCTGCCATTTCTCAGGAGGAGTATAC 5598  
QY 5581 TTGCGTCTCTCATACCAAGGATAAATTAAAGGTCAATTTGCTGCAAGGAAACCTCGTGAAG 5640  
Db 5599 TTGCGTCTCTCATACCAAGGATAAATTAAAGGTCAATTTGCTGCAAGGAAACCTCGTGAAG 5658  
QY 5641 GAGTCCGCGCATGAACACCAACCCGCGCCGCTCCACCTCCCGCAGCAGCAGCCCAACAGCGA 5700  
Db 5659 GAGTCCGCGCATGAACACCAACCCGCGCCGCTCCACCTCCCGCAGCAGCAGCCCAACAGCGA 5718  
QY 5701 GGCCACCCACGTACAAACGAGCAGCATCAACAGCGCTGCGCTCCAGCCAGCAGCCCGCC 5760  
Db 5719 GGCCACCCACGTACAAACGAGCAGCATCAACAGCGCTGCGCTCCAGCCAGCAGCCCGCC 5778  
QY 5761 GAAGCCCCACGACCCCGCAGAGCCAGCAGCAGCCACCGCTACCGCAGGAGGCGGACC 5820  
Db 5779 GAAGCCCCACGACCCCGCAGAGCCAGCAGCAGCCACCGCTACCGCAGGAGGCGGACC 5838  
QY 5821 GAGTGCAGGAGCAAGTCTCTGCGCCGCCCTCGAGCGAGAGAGTCCCCCGGCGCG 5880  
Db 5839 GAGTGCAGGAGCAAGTCTCTGCGCCGCCCTCGAGCGAGAGAGTCCCCCGGCGCG 5898  
QY 5881 ATGCTCAGCAGCGCGAGAGAGCGGTCCCGCGGAGGCTGTTTGAAGACAGCAGCAGGGGC 5940  
Db 5899 ATGCTCAGCAGCGCGAGAGAGCGGTCCCGCGGAGGCTGTTTGAAGACAGCAGCAGGGGC 5958  
QY 5941 CGGCTGCTCCGAGCGCGTGAAGACCCCGCTGTCAGGTGAACAGGAGGAGGAGCGAG 6000  
Db 5959 CGGCTGCTCCGAGCGCGTGAAGACCCCGCTGTCAGGTGAACAGGAGGAGGAGCGAG 6018  
QY 6001 AGTGCCTCTCAAGTTTTCAGGTTAACACCTGTCACCTTATGATGGAATAAAGAGCTG 6060  
Db 6019 AGTGCCTCTCAAGTTTTCAGGTTAACACCTGTCACCTTATGATGGAATAAAGAGCTG 6078  
QY 6061 GACAACTGCCAGCTAACTGCTCAGTCAATCATCAGCTGATCGAGAAATCCCG 6120  
Db 6079 GACAACTGCCAGCTAACTGCTCAGTCAATCATCAGCTGATCGAGAAATCCCG 6138  
QY 6121 CAGCAGTTGAAAAGTCTGTTCTGAGAACAGATTTATTCG 6159  
Db 6139 CAGCAGTTGAAAAGTCTGTTCTGAGAACAGATTTATTCG 6177

Qy	1	ATGTTGAAGTTC	AAATATG	GAGCGCGAA	TCCTTTGGAT	CTGGTCTG	CTGTAACCCATT	60
Db	1	ATGTTGAAGTTC	AAATATG	GAGCGCGAA	TCCTTTGGAT	CTGGTCTG	CTGTAACCCATT	60
Qy	61	GCCAGCGGGCC	TCCAGGC	TGAATCTG	TTCTCCAGGG	AAACACAC	CCCTTTATGACTCAA	120
Db	61	GCCAGCGGGCC	TCCAGGC	TGAATCTG	TTCTCCAGGG	AAACAC	CCCTTTATGACTCAA	120
Qy	121	CAGCAGATG	TCCTCTT	TCGAGAAGG	ATATAGAT	GCCTCTTTG	TTCTCTTTGAA	180
Db	121	CAGCAGATG	TCCTCTT	TCGAGAAGG	ATATAGAT	GCCTCTTTG	TTCTCTTTGAA	180
Qy	181	GAATGCAGT	CAGCTGCT	CTCATGA	ATTAAAGC	ACCTGAGCA	ACTTTGTCGGGAAGTAT	240
Db	181	GAATGCAGT	CAGCTGCT	CTCATGA	ATTAAAGC	ACCTGAGCA	ACTTTGTCGGGAAGTAT	240
Qy	241	TCCGACACCA	TAGTCAG	TTACAGG	AGCTCCAGC	CTTCGGCA	AAAGACTTCGGAAGTCAGA	300
Db	241	TCCGACACCA	TAGTCAG	TTACAGG	AGCTCCAGC	CTTCGGCA	AAAGACTTCGGAAGTCAGA	300
Qy	301	AGTCTTG	TAGTTG	TGCTCACT	TTGCTGA	AGTGCAG	TGTTAAGAGAGAAGCAACCGGG	360
Db	301	AGTCTTG	TAGTTG	TGCTCACT	TTGCTGA	AGTGCAG	TGTTAAGAGAGAAGCAACCGGG	360
Qy	361	GACATCTAT	CTATGA	AGTGATGA	AGAAAGC	TTATTGG	CCCGAGGACAGGTTTCA	420
Db	361	GACATCTAT	CTATGA	AGTGATGA	AGAAAGC	TTATTGG	CCCGAGGACAGGTTTCA	420
Qy	421	TTTTTTG	AGAGCGG	ACATATTTAT	CTCGA	GCACAGCCCG	TCGGATCCCCCAATTA	480
Db	421	TTTTTTG	AGAGCGG	ACATATTTAT	CTCGA	GCACAGCCCG	TCGGATCCCCCAATTA	480
Qy	481	CAGTATGCC	TTTCAGGA	CAAAAATCAC	CTTTATCTG	ATGGAGGA	ATATCAGCCTTGAGGG	540
Db	481	CAGTATGCC	TTTCAGGA	CAAAAATCAC	CTTTATCTG	ATGGAGGA	ATATCAGCCTTGAGGG	540
Qy	541	GACTTGCTG	TCAC	TTTTGA	TAGATATG	AGGACCA	GTTAGATGAAACCTTGATACAGTTT	600
Db	541	GACTTGCTG	TCAC	TTTTGA	TAGATATG	AGGACCA	GTTAGATGAAACCTTGATACAGTTT	600
Qy	601	TACCTAGCT	CAGCTG	ATTTGG	CTGTTT	CACAGCG	TTTCACTGTGGATACGTCATCGA	660
Db	601	TACCTAGCT	CAGCTG	ATTTGG	CTGTTT	CACAGCG	TTTCACTGTGGATACGTCATCGA	660
Qy	661	GACATCAAG	CGCTG	AGAACAT	CTCTG	TGACCG	CACAGACACATCAAGCTGGTGGATTTT	720
Db	661	GACATCAAG	CGCTG	AGAACAT	CTCTG	TGACCG	CACAGACACATCAAGCTGGTGGATTTT	720
Qy	721	GGATCTGCC	CGGAAAT	TGAATTC	AAACA	GATG	TGTAATGCCAACTCCCGATTTGGGACC	780
Db	721	GGATCTGCC	CGGAAAT	TGAATTC	AAACA	GATG	TGTAATGCCAACTCCCGATTTGGGACC	780
Qy	781	CCGATTAC	ATGCTG	CTGAA	GTCTG	ACTCTG	ATGAACGGGGATGAAAAAGGCACCTTAC	840
Db	781	CCGATTAC	ATGCTG	CTGAA	GTCTG	ACTCTG	ATGAACGGGGATGAAAAAGGCACCTTAC	840
Qy	841	GGCCTG	ACTGTG	ACTGG	TGGTGG	TGATTC	CCCTATGAGATGATTTATTTGGGAGA	900
Db	841	GGCCTG	ACTGTG	ACTGG	TGGTGG	TGATTC	CCCTATGAGATGATTTATTTGGGAGA	900
Qy	901	TCCCCCT	TCGACAG	GGAAAC	CTCTGCC	AGAAC	CTTTCAATAACATTTTCCACGGG	960
Db	901	TCCCCCT	TCGACAG	GGAAAC	CTCTGCC	AGAAC	CTTTCAATAACATTTTCCACGGG	960
Qy	961	TTTTTGA	ATTTTC	AGATG	ATCCCC	AAAGT	GAGCAGTACTTTCTTGATCTGATTTCAAAGC	1020
Db	961	TTTTTGA	ATTTTC	AGATG	ATCCCC	AAAGT	GAGCAGTACTTTCTTGATCTGATTTCAAAGC	1020
Qy	1021	TTGTTG	TCGGC	CAGAAAG	AGAGCTCA	AGTTTGA	AGGTCTTTGCTGCCATCTCTTCTTC	1080
Db	1021	TTGTTG	TCGGC	CAGAAAG	AGAGCTCA	AGTTTGA	AGGTCTTTGCTGCCATCTCTTCTTC	1080
Qy	1081	TCTAAAA	ATTG	ACTGG	AAACAC	ATTCG	TACTCTCCCCCTTCGTTTCCACCCCTCAAG	1140

[illegible]



QY 3253 TGTGCGGTTGAGAGCTGAGAGATGCTGGACACCGAGAAAACAGAGCAGGCGGAGAGCC 3312  
Db 3301 TGTGCGGTTGAGAGCTGAGAGATGCTGGACACCGAGAAAACAGAGCAGGCGGAGAGCC 3360  
QY 3313 GATCAGCGATCACCGAGTCTCGCAGGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAAAGCT 3372  
Db 3361 GATCAGCGATCACCGAGTCTGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAAAGCT 3420  
QY 3373 GAGATTCTCGTCTGAGCAGGCTCTCAAAGAGCAGAAAGCTGAAAGCGGAGAGCTCTCT 3432  
Db 3421 GAGATTCTCGTCTGAGCAGGCTCTCAAAGAGCAGAAAGCTGAAAGCGGAGAGCTCTCT 3480  
QY 3433 GACAGCTCAATGACTGAGAGAAAGCATGCTATGCTTTGAAATGAAATGCCCGAAGCTTA 3492  
Db 3481 GACAGCTCAATGACTGAGAGAAAGCATGCTATGCTTTGAAATGAAATGCCCGAAGCTTA 3540  
QY 3493 CAGCAGAAGCTGAGACTGAAAGAGCTCAACAGAGGCTCTGGAAGAGCAAGCCAAA 3552  
Db 3541 CAGCAGAAGCTGAGACTGAAAGAGCTCAACAGAGGCTCTGGAAGAGCAAGCCAAA 3600  
QY 3553 TTACAGCAGAGTGAACCTGAGAAAATCACATTTCCGTCTGACTCAAGGACTGCA 3612  
Db 3601 TTACAGCAGAGTGAACCTGAGAAAATCACATTTCCGTCTGACTCAAGGACTGCA 3660  
QY 3613 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAAGTGAAGTCTGAGTATCAGCTG 3672  
Db 3661 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAAGTGAAGTCTGAGTATCAGCTG 3720  
QY 3673 GAAAACATTCAGTTCTTATCTCTGTAAGAAAGTGAAGTGAAGGACATTTCTCAA 3732  
Db 3721 GAAAACATTCAGTTCTTATCTCTGTAAGAAAGTGAAGTGAAGGACATTTCTCAA 3780  
QY 3733 CAAACCAACTCATGATTTCTGCAAGCAAAATGACCAACTGCTTAAAGAAAAG 3792  
Db 3781 CAAACCAACTCATGATTTCTGCAAGCAAAATGACCAACTGCTTAAAGAAAAG 3838  
QY 3793 GGTTTATTTAGTCAGCGAAAGGACCCTGCTTTACCCACACAGGTTCTCTGAGTAC 3852  
Db 3839 -----AGTTCTCTGAGTAC 3855  
QY 3853 AATGAGCTGAAGTGGCCCTGAGAGGAGAAAGCTCGTGTGCAGAGCTAGAGAGCC 3912  
Db 3856 AATGAGCTGAAGTGGCCCTGAGAGGAGAAAGCTCGTGTGCAGAGCTAGAGAGCC 3915  
QY 3913 CTTTCAAGAACCGCATCGAGTCCGGTCCGCCCGGAGGAAAGCTGCCACCGCAAGCA 3972  
Db 3916 CTTTCAAGAACCGCATCGAGTCCGGTCCGCCCGGAGGAAAGCTGCCACCGCAAGCA 3975  
QY 3973 ACGGACACCCACACCATCCAGCCAGCCACCGGAGGAGAGATGCCCATGTCCGCC 4032  
Db 3976 ACGGACACCCACACCATCCAGCCAGCCACCGGAGGAGAGATGCCCATGTCCGCC 4035  
QY 4033 ATCGTGGGTGCGCAGAGCACAGCCAGTGCATGAGCTGTGSCCCCGCATCCAGC 4092  
Db 4036 ATCGTGGGTGCGCAGAGCACAGCCAGTGCATGAGCTGTGSCCCCGCATCCAGC 4095  
QY 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTCGCGCTTTAAGGAAACGCAATGCAC 4152  
Db 4096 CGCAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTCGCGCTTTAAGGAAACGCAATGCAC 4155  
QY 4153 CACAAATATTCCTCACCGATTCACGTAGACTGAAATGCGAGCCCAAGTGTGCTGTG 4212  
Db 4156 CACAAATATTCCTCACCGATTCACGTAGACTGAAATGCGAGCCCAAGTGTGCTGTG 4215  
QY 4213 TGTCTGGATACCGTGCATTTGAGCCCGCAGGATCCAAATGCTCGAAATGTCAGGTGATG 4272  
Db 4216 TGTCTGGATACCGTGCATTTGAGCCCGCAGGATCCAAATGCTCGAAATGTCAGGTGATG 4275  
QY 4273 TGTCAACCCCAAGTGTCCAGTGTCTGCGAGCCACCTGGCGTGTGCTGCTGAATATGCC 4332  
Db 4276 TGTCAACCCCAAGTGTCCAGTGTCTGCGAGCCACCTGGCGTGTGCTGCTGAATATGCC 4335

QY 4333 ACACACTTCACCGAGGCTTTCTGCGGTGACAAATGAACCTCCAGGTCTCCAGACCAAG 4392  
Db 4336 ACACACTTCACCGAGGCTTTCTGCGGTGACAAATGAACCTCCAGGTCTCCAGACCAAG 4395  
QY 4393 GAGCCCGAGCAGCTTTCACCTTGGAAAGGTGGATGAAGTGGCCAGGAATAACAAACGA 4452  
Db 4396 GAGCCCGAGCAGCTTTCACCTTGGAAAGGTGGATGAAGTGGCCAGGAATAACAAACGA 4455  
QY 4453 GGACAGCAAGGCTGGACAGGAAGTACATTTGCTTGGAGGATCAAAAGTCTCATTTAT 4512  
Db 4456 GGACAGCAAGGCTGGACAGGAAGTACATTTGCTTGGAGGATCAAAAGTCTCATTTAT 4515  
QY 4513 GACAAATGAAGCCAGAGAAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4572  
Db 4516 GACAAATGAAGCCAGAGAAGCTGGACAGAGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4575  
QY 4573 GACGGGATGTATCTATTCAATGTCGCTTGGTCTTCCGAACTCGGAAATACAGCCAAA 4632  
Db 4576 GACGGGATGTATCTATTCAATGTCGCTTGGTCTTCCGAACTCGGAAATACAGCCAAA 4635  
QY 4633 GCA----- 4635  
Db 4636 GCAGATGCCCATACATCTGAAGATGGAATCTCACCGGCACACCACTGCTGCTGCCCGGG 4695  
QY 4636 ----- 4635  
Db 4696 AGAACCTTACTTGTCTAGTCCAGCTTCCCTGACAAACAGCGCTGGTCAACGCCCTTA 4755  
QY 4636 -----GAAAAGCCAGAGCTGATGCTTAAATCTG 4662  
Db 4756 GAATCAGTGTGCGAGGTGGAGAGTTTTCAGGAAAAGCAGAAAGCTGCTTAAATCTG 4815  
QY 4663 CTTTGGAAACTCCCTGCTGAAACTCGAAGGTGATGACCGTCTAGACATGAATGCAACGCTG 4722  
Db 4816 CTTTGGAAACTCCCTGCTGAAACTCGAAGGTGATGACCGTCTAGACATGAATGCAACGCTG 4875  
QY 4723 CCTTTCAGTGACAGGTGGTGTGGTGGGCACCGAGGAAGGGCTCTAGCCCTGAAATGTC 4782  
Db 4876 CCTTTCAGTGACAGGTGGTGTGGTGGGCACCGAGGAAGGGCTCTAGCCCTGAAATGTC 4935  
QY 4783 TTGAAAACCTCCCTAAACCATGTCCAGGAATTTGAGCAGTCTTCCAAATTTATATTATC 4842  
Db 4936 TTGAAAACCTCCCTAAACCATGTCCAGGAATTTGAGCAGTCTTCCAAATTTATATTATC 4995  
QY 4843 AAGAACCTTGGAGAAAGCTACTCATATAGCAGGAGAGAGAGCGGACATGTGTCTTGTGAC 4902  
Db 4996 AAGAACCTTGGAGAAAGCTACTCATATAGCAGGAGAGAGAGCGGACATGTGTCTTGTGAC 5055  
QY 4903 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4962  
Db 5056 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCCAGCCGACATCTCA 5115  
QY 4963 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGGCAGGCAAGATTTGAGAAC 5022  
Db 5116 CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGGCAGGCAAGATTTGAGAAC 5175  
QY 5023 GGGCTCTGCATCTGTGAGCCATGCCAGCAAGTCTGCTTCTCCGCTACAAAGAAAC 5082  
Db 5176 GGGCTCTGCATCTGTGAGCCATGCCAGCAAGTCTGCTTCTCCGCTACAAAGAAAC 5235  
QY 5083 CTCAGCAAAATCTGCATCCGAAAAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5142  
Db 5236 CTCAGCAAAATCTGCATCCGAAAAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC 5295  
QY 5143 TTCAACCAATTCAGTATCTCATTTGGAACCAATAAATTTACGAAATCGACATGAAGAG 5202  
Db 5296 TTCAACCAATTCAGTATCTCATTTGGAACCAATAAATTTACGAAATCGACATGAAGAG 5355  
QY 5203 TACAGCTCGAGGAATTCCTGGATAAGATGACCATTTCTTGGCACCTGCTGCTTTGCC 5262  
Db 5356 TACAGCTCGAGGAATTCCTGGATAAGATGACCATTTCTTGGCACCTGCTGCTTTGCC 5415  
QY 5263 GCCTCTTCCAAACAGCTTTCCCTGTCTCAATCTGAGGTGAACAGCGCAGGCGAGAG 5322

```
Db 5416 GCCTCTTCCAAAGCTTCCCTGTCTCTAATCGTGAAGTGAACGCGAGGCGAGGAG 5475
Qy 5323 GAGTACTTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAG 5382
Db 5476 GAGTACTTGTGTGTTTCCAGAAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAG 5535
Qy 5383 CGACACAGCATCTCAAGTGGAGTCGTCTTACCTTTGGCCCTTTCCTTACAGAGAACCTTAT 5442
Db 5536 CGACACAGCATCTCAAGTGGAGTCGTCTTACCTTTGGCCCTTTCCTTACAGAGAACCTTAT 5595
Qy 5443 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCTCTCAGCA 5502
Db 5596 CTGTTTGTGACCCACTTCAACTCACTCGAAGTAATTTGAGATCCAGGACGCTCTCTCAGCA 5655
Qy 5503 GGAACCCCTGCCGAGGCTTACCTGGACATCCCGAACCCCGGCTTACCTGGGCCCTGGCATTT 5562
Db 5656 GGAACCCCTGCCGAGGCTTACCTGGACATCCCGAACCCCGGCTTACCTGGGCCCTGGCATTT 5715
Qy 5563 TCCTCAGGAGGATTTACTTTGGCGTCTCATACAGGATAAATTAAGGGTTCATTTCTGTCG 5622
Db 5716 TCCTCAGGAGGATTTACTTTGGCGTCTCATACAGGATAAATTAAGGGTTCATTTCTGTCG 5775
Qy 5623 AAGGAAACCTGTGAAGAGTCCGGACCTGAACACACCGGGGCCCGTCCACCTCCCGC 5682
Db 5776 AAGGAAACCTGTGAAGAGTCCGGACCTGAACACACCGGGGCCCGTCCACCTCCCGC 5835
Qy 5683 AGCAGCCCCAACAGGAGGAGCCACCCACGTACAGGACACATCACCAAGGCGGTGGCC 5742
Db 5836 AGCAGCCCCAACAGGAGGAGCCACCCACGTACAGGACACATCACCAAGGCGGTGGCC 5895
Qy 5743 TCCAGCCAGCGCCGCCGGAAGGCCCGCAGCCACCGCGAGAGCCAAAGCACACCCACCCG 5802
Db 5896 TCCAGCCAGCGCCGCCGGAAGGCCCGCAGCCACCGCGAGAGCCAAAGCACACCCACCCG 5955
Qy 5803 TACCGGAGGGGGCGGACCGAGTCCGAGGGAAGTCTCTTGGCCGCGCCCTCTGGAGCA 5862
Db 5956 TACCGGAGGGGGCGGACCGAGTCCGAGGGAAGTCTCTTGGCCGCGCCCTCTGGAGCA 6015
Qy 5863 GAGAAGTCCCGCGCGGATGCTCAGCACGCGGAGAGCGGTCCCGGGAGGCTGTTT 5922
Db 6016 GAGAAGTCCCGCGCGGATGCTCAGCACGCGGAGAGCGGTCCCGGGAGGCTGTTT 6075
Qy 5923 GAGACAGCAGCAGGGGGCGGCTGCTGCGGAGCCGTGAGGACCCCGCTGTCCAGGTTG 5982
Db 6076 GAGACAGCAGCAGGGGGCGGCTGCTGCGGAGCCGTGAGGACCCCGCTGTCCAGGTTG 6135
Qy 5983 AACAAGGGAAGGCGAGAGTGCCTCTCAAGTTTTCACGGTTAAACACTGTCACTATTAT 6042
Db 6136 AACAAGTCTGGGACCACTTTCAGTATTAATCTCAGCCAGAAAACCAACTCCTCATCT 6195
```

## RESULT 3

AAL55217

ID AAL55217 standard; DNA; 6156 BP.

XX

AC AAL55217;

XX

DT 01-MAY-2003 (first entry)

XX

DE Human CR1K related DNA sequence, SEQ ID No 8.

XX

KW Anorectic; hypotensive; cardiac; antilipemic; cerebroprotective;

KW antitumor; osteopathic; antiarthritic; cytosolic; antidepressant;

KW immunomodulator; antihuman; tranquilliser; antiparkinsonian; nootropic;

KW neuroprotective; antiinflammatory; antidiabetic; analgesic;

KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;

KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;

KW central nervous system disorder; chronic obstructive pulmonary disease;

KW diabetes; pain; ds.

XX

OS Homo sapiens.

XX

```
PN WO2003004523-A1.
XX 16-JAN-2003.
XX 28-JUN-2002; 2002WO-EP007156.
XX 02-JUL-2001; 2001US-0301841P.
PR 11-DEC-2001; 2001US-0338651P.
PR 25-APR-2002; 2002US-0375014P.
XX (FARB ) BAYER AG.
XX Zhu Z;
XX WPI; 2003-221576/21.
DR New human citron rho/rac-interacting kinase (CR1K) polypeptide and
PT polynucleotide, useful in preventing, ameliorating or treating diseases
PT associated with human CR1K dysfunction, e.g. obesity, diabetes or
PT Alzheimer's disease.
XX Disclosure; Page 217-222; 237pp; English.
XX The invention relates to an isolated polynucleotide encoding a human
CC citron rho/rac-interacting kinase polypeptide. The isolated
CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the
CC specification. The human citron rho/rac-interacting kinase (CR1K)
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or
CC treating diseases associated with human CR1K dysfunction such as obesity
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery
CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of
CC cancer including endometrial, breast, prostate and colon cancer),
CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's
CC disease), chronic obstructive pulmonary disease, or diabetes. These can
CC also be used to treat pain associated with the disorders. The human CR1K
CC polypeptide is also useful in diagnostic assays or in genetic testing.
CC The expression vector or the reagent is useful in preparing a medicament
CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a
CC central nervous system disorder, or chronic obstructive pulmonary
CC disease. The fusion protein is useful for generating antibodies against a
CC CR1K polypeptide and for use in various assay systems. The methods are
CC useful in producing and detecting the polynucleotide and polypeptide and
CC in screening for agents that modulate the activity of the human CR1K
CC polypeptide. This polynucleotide sequence represents a DNA sequence
CC relating to the human CR1K protein of the invention
XX
SQ Sequence 6156 BP; 1732 A; 1548 C; 1679 G; 1197 T; 0 U; 0 Other;
```

Query Match 92.0%; Score 5666.6; DB 8; Length 6156;

Best Local Similarity 95.9%; Pred. No. 0;

Matches 5945; Conservative 0; Mismatches 9; Indels 243; Gaps 3;

```
Qy 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTGGATGCTGCTGTGTAACCCATT 60
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTGGATGCTGCTGTGTAACCCATT 60
Qy 61 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120
Db 61 GCCAGCGCGGCTCCAGGCTGAATCTCTTCCAGGGGAAACACCCCTTTATGACTCAA 120
Qy 121 CAGCAGATGTCCTCTTTCCGAGAGGGATATTAGATCCCTCTTTGTCTCTTTGAA 180
Db 121 CAGCAGATGTCCTCTTTCCGAGAGGGATATTAGATCCCTCTTTGTCTCTTTGAA 180
Qy 181 GAATGCAAGTCAAGTCTCTTCCGAGAGGGATATTAGATCCCTCTTTGTCTCTTTGAA 240
Db 181 GAATGCAAGTCAAGTCTCTTCCGAGAGGGATATTAGATCCCTCTTTGTCTCTTTGAA 240
Qy 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 300
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 300
```

QY 301 AGCTTTGTAGGTTGGTCACTTTGCTGAAGTGCAGGTGGTAAGAGAGAAAACCAACCGGG 360  
Db 301 AGCTTTGTAGGTTGGTCACTTTGCTGAAGTGCAGGTGGTAAGAGAGAAAACCAACCGGG 360  
QY 361 GACATCTATGCTATGAAAGTGATGAAGAGAGAGAGGCTTTATTTGGCCCGAGGAGAGGTTTCA 420  
Db 361 GACATCTATGCTATGAAAGTGATGAAGAGAGAGAGGCTTTATTTGGCCCGAGGAGAGGTTTCA 420  
QY 421 TTTTGTGAGAGAGCGGACATATTTATCTCAAGCACAAAGCCCGTGGATCCCCCAATTA 480  
Db 421 TTTTGTGAGAGAGCGGACATATTTATCTCAAGCACAAAGCCCGTGGATCCCCCAATTA 480  
QY 481 CAGTATGCCCTTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540  
Db 481 CAGTATGCCCTTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCCTGGAGGG 540  
QY 541 GACTTGCTGTCACTTTTGAATAGATGATGAGGACCAAGTTAGATGAAAACTGATACAGTTT 600  
Db 541 GACTTGCTGTCACTTTTGAATAGATGATGAGGACCAAGTTAGATGAAAACTGATACAGTTT 600  
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTACACGCTTCACTGATGGGATACGTCATCGA 660  
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTACACGCTTCACTGATGGGATACGTCATCGA 660  
QY 661 GACATCAAGCCTGAGAACATTTCTCGTGTACCCGACAGACACATCAAGCTGGTGAATTT 720  
Db 661 GACATCAAGCCTGAGAACATTTCTCGTGTACCCGACAGACACATCAAGCTGGTGAATTT 720  
QY 721 GGATCTGCCCGGAAATGAATTTCAACAGAGTGGTGAATGCCAACTCCCGATTGGGACC 780  
Db 721 GGATCTGCCCGGAAATGAATTTCAACAGAGTGGTGAATGCCAACTCCCGATTGGGACC 780  
QY 781 CCAGATTACATGGCTCCTGAAGTGTGATGAGGAGGATGAAAGGACACCTAC 840  
Db 781 CCAGATTACATGGCTCCTGAAGTGTGATGAGGAGGATGAAAGGACACCTAC 840  
QY 841 GGCCTGGAGTGTGACTGGTGGTCACTGGGCGGTGATTCCTATGAGATGATTTATGGGAGA 900  
Db 841 GGCCTGGAGTGTGACTGGTGGTCACTGGGCGGTGATTCCTATGAGATGATTTATGGGAGA 900  
QY 901 TCCGCTTTCGAGAGGGAACCTTCTGCCAGAACCTTTCAATAACATPATGAATTCAGCGG 960  
Db 901 TCCGCTTTCGAGAGGGAACCTTCTGCCAGAACCTTTCAATAACATPATGAATTCAGCGG 960  
QY 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAGCAGTGAATTTCTGTGATCTGATCAAAAGC 1020  
Db 961 TTTTGTGAATTTCCAGATGACCCCAAGTGAGCAGTGAATTTCTGTGATCTGATCAAAAGC 1020  
QY 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGTGCGCATCCTTTCTTC 1080  
Db 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGGTCTTTGTGCGCATCCTTTCTTC 1080  
QY 1081 TCTAAATTTGACTGAAACACATTTGTTAACTCTCTCCGCCCTTGGTTCCACCCCTCAAG 1140  
Db 1081 TCTAAATTTGACTGAAACACATTTGTTAACTCTCTCCGCCCTTGGTTCCACCCCTCAAG 1140  
QY 1141 TCTGACGATGACACCTCCAAATTTGATGAACACAGAGAGAAATTCGTGGGTTTCATCCCT 1200  
Db 1141 TCTGACGATGACACCTCCAAATTTGATGAACACAGAGAGAAATTCGTGGGTTTCATCCCT 1200  
QY 1201 CCGTGCAGCTGAGCCCTTCAGGCTTCTGGGTGAGAACTGCCCTTTGTGGGTTTTTCG 1260  
Db 1201 CCGTGCAGCTGAGCCCTTCAGGCTTCTGGGTGAGAACTGCCCTTTGTGGGTTTTTCG 1260  
QY 1261 TACAGCAGGCACTGGGATTCCTGGTAGATCTGAGTCTGTTGTCGCGGTCTGACCTCC 1320  
Db 1261 TACAGCAGGCACTGGGATTCCTGGTAGATCTGAGTCTGTTGTCGCGGTCTGACCTCC 1320  
QY 1321 CCGTGCAGCTGAGCTCCATGGAAGAACTTTCTCATCAAAAGCAAGAGACTACAAGAC 1380  
Db 1321 CCGTGCAGCTGAGCTCCATGGAAGAACTTTCTCATCAAAAGCAAGAGACTACAAGAC 1380  
QY 1381 TCTCAGGACAGTGTACAGAGATGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1440

Db 1381 TCTCAGGACAGTGTACAGAGATGAGCAGGAAATGACCCGGTTACATCGAGAGTGTCA 1440  
QY 1441 GAGGTGAGGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500  
Db 1441 GAGGTGAGGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500  
QY 1501 TCCCTCTCTGAGCAGGACCTTGTCTACCTACATCACAGAAATGCAGTAGCTTAAAGCGAAGT 1560  
Db 1501 TCCCTCTCTGAGCAGGACCTTGTCTACCTACATCACAGAAATGCAGTAGCTTAAAGCGAAGT 1560  
QY 1561 TTGAGCAAGCAGCGATGGAGGTGTCCAGAGGATGACAAAGCACTGACGCTTCTCCAT 1620  
Db 1561 TTGAGCAAGCAGCGATGGAGGTGTCCAGAGGATGACAAAGCACTGACGCTTCTCCAT 1620  
QY 1621 GATATCAGAGCAGCAGCGGAGAGCTCCAAAGAAATCAAAGCAGGAGTAGTACCAGGCTCAA 1680  
Db 1621 GATATCAGAGCAGCAGCGGAGAGCTCCAAAGAAATCAAAGCAGGAGTAGTACCAGGCTCAA 1680  
QY 1681 GTGGAAGAAATGAGGTTGATGATGAATCACTTGTGAAGAGGATCTTGTCTCAGCAAGAAGA 1740  
Db 1681 GTGGAAGAAATGAGGTTGATGATGAATCACTTGTGAAGAGGATCTTGTCTCAGCAAGAAGA 1740  
QY 1741 CGGAGTCACTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1800  
Db 1741 CGGAGTCACTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1800  
QY 1801 CGGAAACCGCAGAAATGTCAAGATAAACTTTGAAGGCTAAGGATCAAGGAGAGCCTGAA 1860  
Db 1801 CGGAAACCGCAGAAATGTCAAGATAAACTTTGAAGGCTAAGGATCAAGGAGAGCCTGAA 1860  
QY 1861 GTGGGAAATATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
Db 1861 GTGGGAAATATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGGAG 1920  
QY 1921 CTCACAGAGAAACTGGAGAAAGGCT----- 1944  
Db 1921 CTCACAGAGAAACTGGAGAAAGGCTGTAAAGCCAGCAGCGAGGCCACCGAGCTGCTGCAG 1980  
QY 1945 -----GCAAGAGAGCGAGCGAGAGGAGCTGGAGAAAGCTTGAGAACCGAGAG 1992  
Db 1981 AATATCGCGCAGGCAAGAGGAGCGAGCGAGGAGGAGCTGGAGAAAGCTTGAGAACCGAGAG 2040  
QY 1993 GATTCTTCTCAAGGCTACAGAAAGAGCTGCTGGAAGCTGAGGAAAGCGCCCACTTCTCTG 2052  
Db 2041 GATTCTTCTCAAGGCTACAGAAAGAGCTGCTGGAAGCTGAGGAAAGCGCCCACTTCTCTG 2100  
QY 2053 GAGAAACAGGTAAGAGACTAGAGACCATGGACGCTAGAGAAACAGACTGAAGAGATGAC 2112  
Db 2101 GAGAAACAGGTAAGAGACTAGAGACCATGGACGCTAGAGAAACAGACTGAAGAGATGAC 2160  
QY 2113 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATTAATAATTTGAGGCTCGAGAG 2172  
Db 2161 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATTAATAATTTGAGGCTCGAGAG 2220  
QY 2173 AAACATCGGAGGCGCCAAAGTCTCAGCCCACTAGAAAGTGCACCTGAAAACAGAAAGAG 2232  
Db 2221 AAACATCGGAGGCGCCAAAGTCTCAGCCCACTAGAAAGTGCACCTGAAAACAGAAAGAG 2280  
QY 2233 CAGCACTATCAGGAAAGATTAAGTTTGGCAAAATCAGATAAAGAAAGCCTGGCTGAC 2292  
Db 2281 CAGCACTATCAGGAAAGATTAAGTTTGGCAAAATCAGATAAAGAAAGCCTGGCTGAC 2340  
QY 2293 AAGGAGACCTGAGAAACATGATGAGAGACAGAGAGGAGGCCCATGAGAAAGGCAAA 2352  
Db 2341 AAGGAGACCTGAGAAACATGATGAGAGACAGAGAGGAGGCCCATGAGAAAGGCAAA 2400  
QY 2353 ATTCTCAGCAACAGAGGCGATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAA 2412  
Db 2401 ATTCTCAGCAACAGAGGCGATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAA 2460  
QY 2413 CAGAGGATTTGTGAACTGTCTGAGGCCAATAACTTTGACGCAAAATAGCAGTCTTTTACC 2472

Db 2461 CAGAGGATTGTGAACTGTCTGAAGCCAAATAAACTTGCAGCAAAATAGCAGTCTTTTACC 2520  
Qy 2473 CAAAGGAAATGAAGCCCAAGAGAGATGATTCTGAACTCAGGCAACAGAAAAATTTTAC 2532  
Db 2521 CAAAGGAAATGAAGCCCAAGAGAGATGATTCTGAACTCAGGCAACAGAAAAATTTTAC 2580  
Qy 2533 CTGGAGACAGGCTGGGAAGTTGGAGGCCAGAACCGGAAATCTGGAGGAGCAGCTGGAG 2592  
Db 2581 CTGGAGACAGGCTGGGAAGTTGGAGGCCAGAACCGGAAATCTGGAGGAGCAGCTGGAG 2640  
Qy 2593 AAGATCAGCCACCAAGACACAGTGAAGAATCGGCTCTGAACTGGAGCAAGATTG 2652  
Db 2641 AAGATCAGCCACCAAGACACAGTGAAGAATCGGCTCTGAACTGGAGCAAGATTG 2700  
Qy 2653 CGGAGGTCAGTCTAGAGCAGGAGGAGAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2712  
Db 2701 CGGAGGTCAGTCTAGAGCAGGAGGAGAGAACTGGAGCTCAAGCGCCAGCTCAGAG 2760  
Qy 2713 CTAAGCTCTCCCTGAGAGCGGAGTCACAGTTGAGAGCCCTGAGGCTGACGGGG 2772  
Db 2761 CTAAGCTCTCCCTGAGAGCGGAGTCACAGTTGAGAGCCCTGAGGCTGACGGGG 2820  
Qy 2773 GCCTCGAGAGCAGCTTCGCGAGGCGAAGACAGAGCTGGAAGAGACACACAGCAAGCT 2832  
Db 2821 GCCTCGAGAGCAGCTTCGCGAGGCGAAGACAGAGCTGGAAGAGACACACAGCAAGCT 2880  
Qy 2833 GAAGAGGATCCAGGCACTCAGGCACATAGAGATGAATCCAGGCGMAATTTGATGCT 2892  
Db 2881 GAAGAGGATCCAGGCACTCAGGCACATAGAGATGAATCCAGGCGMAATTTGATGCT 2940  
Qy 2893 CTTTGTAAAGCTGTACTGTAACTCAGAGCTGGAGGAGCTAAACAGCTGACCGAG 2952  
Db 2941 CTTTGTAAAGCTGTACTGTAACTCAGAGCTGGAGGAGCTAAACAGCTGACCGAG 3000  
Qy 2953 GAAACGCTGAATCAACCAACCAAACTTCTACTTGTCCAAACAACTCGATGAGCTTCT 3012  
Db 3001 GAAACGCTGAATCAACCAACCAAACTTCTACTTGTCCAAACAACTCGATGAGCTTCT 3060  
Qy 3013 GGGGCCAAGCAGAGATCTGAACCTGCAAGTGAAGTGGAGCCATCTCCGCCGGAGATC 3072  
Db 3061 GGGGCCAAGCAGAGATCTGAACCTGCAAGTGAAGTGGAGCCATCTCCGCCGGAGATC 3120  
Qy 3073 ACGGAACGAGATGAGCTTACAGCCAGAGAACCAAGATGAGGCTCTGAAGACCAAG 3132  
Db 3121 ACGGAACGAGATGAGCTTACAGCCAGAGAACCAAGATGAGGCTCTGAAGACCAAG 3180  
Qy 3133 TGACCATGCTGAGAGAACAGCTCATGATTTGGAGCCCTAAACGATGAGTGTAGAA 3192  
Db 3181 TGACCATGCTGAGAGAACAGCTCATGATTTGGAGCCCTAAACGATGAGTGTAGAA 3240  
Qy 3193 AAAGAGCGGAGTGGAGGCTTGGAGAGCGTCTGGGTGATGAGAAATCCAGATTTGAG 3252  
Db 3241 AAAGAGCGGAGTGGAGGCTTGGAGAGCGTCTGGGTGATGAGAAATCCAGATTTGAG 3300  
Qy 3253 TGTTCGGGTTGAGAGCTGAGAGATGCTGGACACCGAGAAACAGAGAGGCGGAGAGCC 3312  
Db 3301 TGTTCGGGTTGAGAGCTGAGAGATGCTGGACACCGAGAAACAGAGAGGCGGAGAGCC 3360  
Qy 3313 GATCAGCGGATCACCGATCTCCAGAGTGTGGAGCTGGAGTGAAGGAGCACAAGGCT 3372  
Db 3361 GATCAGCGGATCACCGATCTCCAGAGTGTGGAGCTGGAGTGAAGGAGCACAAGGCT 3420  
Qy 3373 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGAGAGAGCTGAAGGCGGAGAGCTCTCT 3432  
Db 3421 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGAGAGAGCTGAAGGCGGAGAGCTCTCT 3480  
Qy 3433 GAGAGCTCAATGACCTGAGAGAGAGCTATGCTTGAATGAATGCAATGCGCCGAGCTTA 3492  
Db 3481 GAGAGCTCAATGACCTGAGAGAGAGCTATGCTTGAATGAATGCAATGCGCCGAGCTTA 3540  
Qy 3493 CAGCAGAAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAGAGAGCAAGCCAAA 3552  
Db 3541 CAGCAGAAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAGAGAGCAAGCCAAA 3600

Qy 3553 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCGTCTGACTCAAGGACTGCAA 3612  
Db 3601 TTACAGCAGCAGATGGACCTGCAGAAAAATCAATTTTCGTCTGACTCAAGGACTGCAA 3660  
Qy 3613 GAAGCTCTAGATCGGCTGATCTACTGAAGCAGAAAGAGTCACTTGGAGTATCAGCTG 3672  
Db 3661 GAAGCTCTAGATCGGCTGATCTACTGAAGCAGAAAGAGTCACTTGGAGTATCAGCTG 3720  
Qy 3673 GAAAAATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAAATGGAAGCACTATTTCTCAA 3732  
Db 3721 GAAAAATTCAGGTTCTCTATTCTCATGAAAAGGTGAAAAATGGAAGCACTATTTCTCAA 3780  
Qy 3733 CAAACCAAACTCATTGATTTTCTGCAAGCCAAAATGGAACCACTCTAAAAGAAAAAG 3792  
Db 3781 CAAACCAAACTCATTGATTTTCTGCAAGCCAAAATGGAACCACTCTAAAAGAAAA-- 3838  
Qy 3793 GGTTTATTTAGTCGACGGAAGAGGACCTGTCTTTACCCACACAGGTTCTCTGCAGTAC 3852  
Db 3839 -----AGGTTCTCTGCAGTAC 3855  
Qy 3853 AATGAGCTGAAGCTGGCCCTGGAGAAAGGAGAAAGCTCGCTGTGCCAGAGCTAGAGAGCC 3912  
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAAAGGAGAAAGCTCGCTGTGCAGAGCTAGAGAGCC 3915  
Qy 3913 CTTTCAAGAACCCGCACTCGAGCTCCGCTCCGCCGGAGGAAGCTGCCACCGCAAGCA 3972  
Db 3916 CTTTCAAGAACCCGCACTCGAGCTCCGCTCCGCCGGAGGAAGCTGCCACCGCAAGCA 3975  
Qy 3973 ACGGACCAACCAACCCATCCAGCCAGCAACCGCGAGGAGCAGATGTCGCGC 4032  
Db 3976 ACGGACCAACCAACCCATCCAGCCAGCAACCGCGAGGAGCAGATGTCGCGC 4035  
Qy 4033 ATCGTGGCTGCGCAGAGCAGAGCCAGTGCATGAGCTGCTGCCCGCCATCCAGC 4092  
Db 4036 ATCGTGGCTGCGCAGAGCAGAGCCAGTGCATGAGCTGCTGCCCGCCATCCAGC 4095  
Qy 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAAACGATGCAC 4152  
Db 4096 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAAACGATGCAC 4155  
Qy 4153 CACAATTTCTCAGCGATTCAAGCTAGAGTGAATGAAATGCGAGCCCAAGAGTGTGCTG 4212  
Db 4156 CACAATTTCTCAGCGATTCAAGCTAGAGTGAATGAAATGCGAGCCCAAGAGTGTGCTG 4215  
Qy 4213 TGTCTGATACCGTGCATTTGAGCGCCAGGCTTCAAAATGCTCGAAATGCTCAGAGTATG 4272  
Db 4216 TGTCTGATACCGTGCATTTGAGCGCCAGGCTTCAAAATGCTCGAAATGCTCAGAGTATG 4275  
Qy 4273 TGTACCCCCAAGTGTCCAGCTGTCAGCCAGCCACCTGCGGCTTGCCTGTGATATGCC 4332  
Db 4276 TGTACCCCCAAGTGTCCAGCTGTCAGCCAGCCACCTGCGGCTTGCCTGTGATATGCC 4335  
Qy 4333 ACACATTTCAACGAGGCTTCTGCCGTGACAAAATGAATCCGCCAGTCTCCAGACCAAG 4392  
Db 4336 ACACATTTCAACGAGGCTTCTGCCGTGACAAAATGAATCCGCCAGTCTCCAGACCAAG 4395  
Qy 4393 GAGCCACGAGCAGCTTGCACTTGGAGGGTGGATGAAGTGCAGGGAATACAAACGA 4452  
Db 4396 GAGCCACGAGCAGCTTGCACTTGGAGGGTGGATGAAGTGCAGGGAATACAAACGA 4455  
Qy 4453 GGACAGCAAGCTGGGACAGAGATGATGCTCTGGAGGATCAAAAGTCTCATTTAT 4512  
Db 4456 GGACAGCAAGCTGGGACAGAGATGATGCTCTGGAGGATCAAAAGTCTCATTTAT 4515  
Qy 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCC 4572  
Db 4516 GACAATGAAGCCAGAGAGCTGGACAGAGCCGCTGGAGAAATTTGAGCTGTGCTTCCC 4575  
Qy 4573 GACGGGATGATCTATTATGCTGCTGCTGCTTCCGAACTCCGAATACAGCCAA 4632  
Db 4576 GACGGGATGATCTATTATGCTGCTGCTTCCGAACTCCGAATACAGCCAA 4635

Qy	4633	GCA-----	4635
Db	4636	GCAGATGTCCTCCATACATACCTGAAGATGGAATCTCACCCGCACACACCTGCTGGCCGGG	4695
Qy	4636	-----	4635
Db	4696	AGAACCTCTACTTGTGCTAGCTCCACGCTCCCTGTGACAAACAGGGCTGGGTACCCGCCCTTA	4755
Qy	4636	-----	4662
Db	4756	GAATCAGTTGTCGACGTGGGAGAGTTTCTAGGGAAAAAGCAGAGCTGATGCTTAAACTG	4815
Qy	4663	CTTGGAAACTCCCTGCTGTAACCTGGAAAGTGATGACCGTCTAGACATGAACCTGCACGCTG	4722
Db	4816	CTTGGAAACTCCCTGCTGTAACCTGGAAAGTGATGACCGTCTAGACATGAACCTGCACGCTG	4875
Qy	4723	CCCTTCAGTACACAGGTGTGTGTGTGGGCACCGAGGAAGGCTCTPACGCCCTGAATGTC	4782
Db	4876	CCCTTCAGTACACAGGTGTGTGTGTGGGCACCGAGGAAGGCTCTPACGCCCTGAATGTC	4935
Qy	4783	TTGAAAAACTCCCTAACCCATGTGCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC	4842
Db	4936	TTGAAAAACTCCCTTAACCCATGTGCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC	4995
Qy	4843	AAGACCTCGAGAAGCTACTCATGATAGCAGGAGAAGCGGCACCTGTGTCTTGTGGAC	4902
Db	4996	AAGACCTCGAGAAGCTACTCATGATAGCAGGAGAAGCGGCACCTGTGTCTTGTGGAC	5055
Qy	4903	GTGAAGAAAGTGAAACAGTCCCTGGCCAGTCCACCTGCCTGCCAGCCGACATCTCA	4962
Db	5056	GTGAAGAAAGTGAAACAGTCCCTGGCCAGTCCACCTGCCTGCCAGCCGACATCTCTCA	5115
Qy	4963	CCCAACATTTTGAAGCTGTCAAGGGGTGCCACTTGTTTGGGCGAGCGAAGATTGAGAAC	5022
Db	5116	CCCAACATTTTGAAGCTGTCAAGGGGTGCCACTTGTTTGGGCGAGCGAAGATTGAGAAC	5175
Qy	5023	GGGCTCTGCACTCTGTGAGCGAATGCCCAGCAAGTGTGTCATTTCTCGCTTACACGAAAC	5082
Db	5176	GGGCTCTGCACTCTGTGAGCGAATGCCCAGCAAGTGTGTCATTTCTCGCTTACACGAAAC	5235
Qy	5083	CTCAGCAATACTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5142
Db	5236	CTCAGCAATACTGCATCCGAAAGAGATAGAGACCTCAGAGCCCTGCAGCTGTATCCAC	5295
Qy	5143	TTCAACCAATTACAGTATCTCTCATTTGGAACCAATAAATTTACGAATTCGACATGAAGCAG	5202
Db	5296	TTCAACCAATTACAGTATCTCTCATTTGGAACCAATAAATTTACGAATTCGACATGAAGCAG	5355
Qy	5203	TACAGCTCGAGGAATTCCTGGATAAGATGACCAATTCCTTGGCACCCTGCTGTTGTGCC	5262
Db	5356	TACAGCTCGAGGAATTCCTGGATAAGATGACCAATTCCTTGGCACCCTGCTGTTGTGCC	5415
Qy	5263	GCCTCTTCCAAACAGTTCCCTGTCTCAATCGTGAGGTGAACCGCAGGGCGAGCGAGAG	5322
Db	5416	GCCTCTTCCAAACAGTTCCCTGTCTCAATCGTGAGGTGAACCGCAGGGCGAGCGAGAG	5475
Qy	5323	GAGTACTTGCTGTGTTTCCACGAATTTGAGTGTCTGTGGATTTCTTACGGAAGACGTAGC	5382
Db	5476	GAGTACTTGCTGTGTTTCCACGAATTTGAGTGTCTGTGGATTTCTTACGGAAGACGTAGC	5535
Qy	5383	CGCACAGACGATCTCAAGTGGAGTCCGCTTACCTTTGGCCTTTGCTACAGAGAACCCCTAT	5442
Db	5536	CGCACAGACGATCTCAAGTGGAGTCCGCTTACCTTTGGCCTTTGCTACAGAGAACCCCTAT	5595
Qy	5443	CTGTTTGTGACCCCATCTTCACTCAGAGTAAATTGAGATCCAGGCACGCTCTCTCAGCA	5502
Db	5596	CTGTTTGTGACCCCATCTTCAACTCAGTGAAGTAAATTGAGATCCAGGCACGCTCTCTCAGCA	5655
Qy	5503	GGGACCCCTGCCGAGCGTACTCTGACATCCCGAAACCCCGCTACCTTGGGCCCTGCCATT	5562
Db	5656	GGGACCCCTGCCGAGCGTACTCTGACATCCCGAAACCCCGCTACCTTGGGCCCTGCCATT	5715
Qy	5563	TCCTCAGGAGCGAATTTACTTGGCGTCTCATACACAGGATAAATTAAGGGTCAATTTGCTGC	5622

Db	5716	TCCTCAGGACGATTACCTTGGCGTCTCATACCGAGTAAATTTAAAGGGTCATTTGTCTGC	5776
Qy	5623	AAGGGAACCTCTGTAAGAGTCCGGCACTGAACACACCGGGGCCGCTCCACCTCCCGC	5682
Db	5776	AAGGGAACCTCTGTAAGAGTCCGGCACTGAACACACCGGGGCCGCTCCACCTCCCGC	5835
Qy	5683	AGCAGCCCCAACAGCGAGCGAGGCCCCACCCACGTACAACGAGCACATCACCAAGCGGTGGCC	5742
Db	5836	AGCAGCCCCAACAGCGAGCGAGGCCCCACCCACGTACAACGAGCACATCACCAAGCGGTGGCC	5895
Qy	5743	TCAGCCACAGCGCCCGCCGGAAGCCCGCCAGCCAACCGGAGAGCGACAGCACACCCACCGC	5802
Db	5896	TCCAGCCACAGCGCCCGCCGGAAGCCCGCCAGCCAACCGGAGAGCGACAGCACACCCACCGC	5955
Qy	5803	TACCGCGAGGGCGGACCGAGCTCGCGAGGACAAGTCTCTGGCGCCGCCCTGGAGCGA	5862
Db	5956	TACCGCGAGGGCGGACCGAGCTCGCGAGGACAAGTCTCTGGCGCCGCCCTGGAGCGA	6015
Qy	5863	GAGAAGTCCCCCGCGCCGATGCTCAGCACCGGAGAGAGCGGTCCCGGAGAGGCTGTTT	5922
Db	6016	GAGAAGTCCCCCGCGCCGATGCTCAGCACCGGAGAGAGCGGTCCCGGAGAGGCTGTTT	6075
Qy	5923	GAAGACAGCAGCGAGGGCCCGCTGCCTCGCGGAGCGGTGAGGACCCCGCTGTCCACAGTG	5982
Db	6076	GAAGACAGCAGCGAGGGCCCGCTGCCTCGCGGAGCGGTGAGGACCCCGCTGTCCACAGTG	6135
Qy	5983	AACAAGGAAGAGGGCA	5999
Db	6136	AACAAGGTGAGGCAGCA	6152
RESULT 4			
AD	AAD38864		
ID	AAD38864	standard; cDNA; 6298 BP.	
XX	AC	AAD38864;	
XX	DT	23-SEP-2002 (first entry)	
XX	DE	Human kinase (PKIN)-21 cDNA.	
XX	KW	Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;	
KW	KW	acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;	
KW	KW	asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;	
KW	KW	development; hepatitis; cardiovascular; hypertension; drug screening;	
KW	KW	myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;	
KW	KW	faty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;	
KW	KW	hypercholesterolaemia; obesity; gene therapy; cytostatic; anti-hiv;	
KW	KW	neuroprotective; hepatotropic; hypotensive; cardiant; nephrotropic;	
XX	XX	hyperlipidaemia; enzyme; gene; ss.	
OS	XX	Homo sapiens.	
XX	XX	Location/Qualifiers	
FT	Key	55..6219	
FT	CDS	/*tag= a	
FT	FT	/product= "Human kinase (PKIN)-21"	
XX	XX	WO200233099-A2.	
PN	PN	25-APR-2002.	
PD	PD	20-OCT-2001; 2001WO-US047728.	
XX	XX	20-OCT-2000; 2000US-0242410P.	
XX	XX	27-OCT-2000; 2000US-0244068P.	
PR	PR	03-NOV-2000; 2000US-0245708P.	
PR	PR	09-NOV-2000; 2000US-0247672P.	
PR	PR	16-NOV-2000; 2000US-0249565P.	
PR	PR	22-NOV-2000; 2000US-0252730P.	
PR	PR	01-DEC-2000; 2000US-0250807P.	
XX	XX		

(INCY-) INCYTE GENOMICS INC.

PA Gururajan R, Baughn MR, Wallia NK, Elliott VS, Xu Y, Arvizu C;  
XX Yao MG, Ramkumar J, Ding L, Tang YT, Hafalia AJA, Nguyen DB;  
PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;  
PI Recipon SA, Lu DAM, Borowsky ML, Thornton M, Swarnaker A;  
XX Thangavelu K, Khan FA, Ison CH;  
DR WPI; 2002-454603/48.  
DR P-PSDB; RAE24150.  
XX

PT New human kinase polypeptide, for diagnosing, preventing and treating  
PT cancer, immune system disorders, growth and development disorders,  
XX cardiovascular disorders and lipid disorders.  
PS Claim 5; Page 207-209; 210pp; English.

CC The invention relates human kinases (PKIN) and their corresponding  
CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,  
CC treating and preventing cancer, an immune system disorder (e.g., acquired  
CC immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,  
CC atherosclerosis, multiple sclerosis, psoriasis), disorders affecting  
CC growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),  
CC cardiovascular disorder (e.g., hypertension, myocardial infarction,  
CC Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,  
CC Gaucher's disease, Niemann-Pick's disease, hypercholesterolaemia,  
CC hyperlipidaemia, obesity), and for assessing the effects of exogenous  
CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a  
CC condition or a disease associated with the expression of PKIN in a  
CC biological sample. A composition comprising PKIN or an agonist or  
CC antagonist of PKIN is useful for treating a disease or condition  
CC associated with decreased or increased expression of functional PKIN.  
CC PKIN is useful in a number of drug screening techniques and to analyse  
CC the proteome of a tissue or cell type. PKIN DNA is useful for creating  
CC knockin humanised animals or transgenic animals to model human diseases,  
CC and in somatic or germline gene therapy. The present sequence is human  
CC PKIN cDNA  
XX

SQ Sequence 6298 BP; 1772 A; 1585 C; 1720 G; 1221 T; 0 U; 0 Other;  
Query Match 92.0%; Score 5666.4; DB 6; Length 6298;  
Best Local Similarity 95.5%; Pred. No. 0;  
Matches 5961; Conservative 0; Mismatches 36; Indels 243; Gaps 3;

QY 1 ATGTTGAAGTTCAATATCGAGCGCGGAATCCTTTGGATGCTGGTGTGTAACCCATT 60  
DB 55 ATGTTGAAGTTCAATATCGAGCGCGGAATCCTTTGGATGCTGGTGTGTAACCCATT 114

QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCTTTATGACTCAA 120  
DB 115 GCCAGCGGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCTTTATGACTCAA 174

QY 121 CAGCAGATGCTCTCTCTTTCCGAGAGAGATATTAGATGCGCTTTTGTCTCTTTGAA 180  
DB 175 CAGCAGATGCTCTCTCTTTCCGAGAGAGATATTAGATGCGCTTTTGTCTCTTTGAA 234

QY 181 GAATGCAGTCAGCTGCTGATGAAGATTAAAGACGTGACACTTTGTCGGAAGTAT 240  
DB 235 GAATGCAGTCAGCTGCTGATGAAGATTAAAGACGTGACACTTTGTCGGAAGTAT 294

QY 241 TCCGACACATAGTGTAGTTACAGAGCTCCAGGCTTCGGCAAGGACTTCGAAGTCAGA 300  
DB 295 TCCGACACATAGTGTAGTTACAGAGCTCCAGGCTTCGGCAAGGACTTCGAAGTCAGA 354

QY 301 AGTCTTGTAGTTGGTTCACCTTTGCTCAAGTCAGGTGTTAAGAGAAAGAACCGGG 360  
DB 355 AGTCTTGTAGTTGGTTCACCTTTGCTCAAGTCAGGTGTTAAGAGAAAGAACCGGG 414

QY 361 GACATCTATGCTATGAAGTATGAAGAGAGAGGCTTTATTGGCCCGAGGAGGTTTCA 420  
DB 415 GACATCTATGCTATGAAGTATGAAGAGAGAGGCTTTATTGGCCCGAGGAGGTTTCA 474

QY 421 TTTTGTGAGGAGAGCGGAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 480

DB 475 TTTTGTGAGGAGAGCGGAACATATTATCTCGAAGCACAGCCCGTGGATCCCCCAATTA 534

QY 481 CAGTATGCTTTTCAGGACAAAATACCTTTATCTGTATGGAGGAATATCAGCTGGAGGG 540

DB 535 CAGTATGCTTTTCAGGACAAAATACCTTTATCTGTGTATGGGAATATCAGCTGGAGGG 594

QY 541 GACTTGTCTGTCATCTTTTGAATAGATATAGGACACAGTGTAGATGAAAACCTGATACAGTTT 600

DB 595 GACTTGTCTGTCATCTTTTGAATAGATATAGGACACAGTGTAGATGAAAACCTGATACAGTTT 654

QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGAGCTTCTATCTGTATGGGATAGTGCATCGA 660

DB 655 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGAGCTTCTATCTGTATGGGATAGTGCATCGA 714

QY 661 GACATCAAGCCTGAGAAATCTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720

DB 715 GACATCAAGCCTGAGAAATCTCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 774

QY 721 GGATCTGCGCGGAAATGAATTTCAAAACAAGATGCTGAATGCCAAACTCCGATTTGGACC 780

DB 775 GGAATCTGCGCGGAAATGAATTTCAAAACAAGATGCTGAATGCCAAACTCCGATTTGGACC 834

QY 781 CCAGATTACATGGCTCTCTGAAAGTCTGACTGTGATGAACGGGATGAAAAGGCACCTAC 840

DB 835 CCAGATTACATGGCTCTCTGAAAGTCTGACTGTGATGAACGGGATGAAAAGGCACCTAC 894

QY 841 GGCTGTGACTGTGACTGCTGCTGAGTGGGCTGATTCCTATGAGATGATTTATGGAGA 900

DB 895 GGCTGTGACTGTGACTGCTGCTGAGTGGGCTGATTCCTATGAGATGATTTATGGAGA 954

QY 901 TCCCTCTTCGAGAGGAACTCTGCCAGAACTTCAATAACATTTGATTTTCAGCGG 960

DB 955 TCCCTCTTCGAGAGGAACTCTGCCAGAACTTCAATAACATTTGATTTTCAGCGG 1014

QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGTATCTGATTTCAAAGC 1020

DB 1015 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGTATCTGATTTCAAAGC 1074

QY 1021 TTGTTGTCGCGCAGAGAGAGACTGAAAGTTTGAAGCTTTTGTGCTCCATCTTTCTTC 1080

DB 1075 TTGTTGTCGCGCAGAGAGAGACTGAAAGTTTGAAGCTTTTGTGCTCCATCTTTCTTC 1134

QY 1081 TCTAAAATTGACTGGAAACCAATTCGTAACCTCTCTCCCTCCCTCCCTCCCTCCCTCAAG 1140

DB 1135 TCTAAAATTGACTGGAAACCAATTCGTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1194

QY 1141 TCTGACGATGACACCTTCCAAATTTTGTATGAACAGAGAGAAATTCGTGGGTTTCATCTCT 1200

DB 1195 TCTGACGATGACACCTTCCAAATTTTGTATGAACAGAGAGAAATTCGTGGGTTTCATCTCT 1254

QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTTCG 1260

DB 1255 CCGTGCAGCTGAGCCCTCAGGCTTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTTCG 1314

QY 1261 TACAGAGGACTGGGATTTCTGTTAGATCTGAGTCTGTTGTTGTTGTTGTTGTTGTTGTT 1320

DB 1315 TACAGAGGACTGGGATTTCTGTTAGATCTGAGTCTGTTGTTGTTGTTGTTGTTGTTGTT 1374

QY 1321 CTTGCCAAGACTAGCTTCCATGAAAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1380

DB 1375 CTTGCCAAGACTAGCTTCCATGAAAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1434

QY 1381 TCTCAGGACAAAGTGTCAAAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1440

DB 1435 TCTCAGGACAAAGTGTCAAAAGATGGAGCAGGAAATGACCCGGTTACATCGGAGAGTGTCA 1494

QY 1441 GAGGTGGAGGCTGTGCTTAGTTCAGAGGAGGTTGAGCTGAAAGGCTCTGAGACTCAGAGA 1500

DB 1495 GAGGTGGAGGCTGTGCTTAGTTCAGAGGAGGTTGAGCTGAAAGGCTCTGAGACTCAGAGA 1554

QY 1501 TCCCTCTTCGAGCAGGACTTGTCTACCTACATCAAGAAATGAGTGTAAAGCGAAGT 1560

Db 1555 TCCCTCTCGGACGAGCCTTGCTACCTACATCAGAAATGCAGTAGCTTAAAGCGAGT 1614  
Qy 1561 TTGGAGCAAGCAGCATGGAGTGTCCAGAGGATGACAAAGCATCTCAGCTTCTCCAT 1620  
Db 1615 T\*GGAGCAAGCAGCATGGAGTGTCCAGAGGATGACAAAGCATCTCAGCTTCTCCAT 1674  
Qy 1621 GATATCAGAGACGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGAGTACAGGCTCAA 1680  
Db 1675 GATATCAGAGACGAGCCGGAAGCTCCAAAGAAATCAAAGAGCAGAGTACAGGCTCAA 1734  
Qy 1681 GTGGAAGAAATGAGTGTGATGATGAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAGA 1740  
Db 1735 GTGGAAGAAATGAGTGTGATGATGAATCAGTTTGGAGAGGATCTTGTCTCAGCAAGAGA 1794  
Qy 1741 CGAGTGTATCTACGAATCTGAGCTGAGAGTCTCGGCTTGCTGCTGAGAAATTCAG 1800  
Db 1795 CGAGTGTATCTACGAATCTGAGCTGAGAGTCTCGGCTTGCTGCTGAGAAATTCAG 1854  
Qy 1801 CGGAAAGCGACAGAAATGTGACATAAATGTTTGAAGGCTTAAGGATCAAAGGGAAGCCTCAA 1860  
Db 1855 CGGAAAGCGACAGAAATGTGACATAAATGTTTGAAGGCTTAAGGATCAAAGGGAAGCCTCAA 1914  
Qy 1861 GTGGAGAAATTCGAAATCTGAGAAAGATCAATGCTGAGAGCAGCTCAAAATTCAGAG 1920  
Db 1915 GTGGAGAAATTCGAAATCTGAGAAAGATCAATGCTGAGAGCAGCTCAAAATTCAGAG 1974  
Qy 1921 CTCCAAGGAAATCTGAGAGGCTG----- 1945  
Db 1975 CTCCAAGGAAATCTGAGAGGCTGTAAGGCCAGCAGGAGCCACCGAGCTGCTGCAG 2034  
Qy 1946 -----CAAAGGAGCGAGCCGAGAGGAGCTGGAGAAGCTCAGAAACCGAGAG 1992  
Db 2035 AATATCCGCGCAGGCAAAAGGAGCGAGCCGAGAGGAGCTGGAGAAGCTCGAGAACCGAGAG 2094  
Qy 1993 GATTCTTCTGAGGCAATCAGAAAGAGCTGTGGAGCTGAGAAACCGGCGCATCTCTG 2052  
Db 2095 GATTCTTCTGAGGCAATCAGAAAGAGCTGTGGAGCTGAGAAACCGGCGCATCTCTG 2154  
Qy 2053 GAGAACAGGTAAGAGACTAGAGACCTAGAGCGCTAGAGAAACAGAGCTGAAGATGAC 2112  
Db 2155 GAGAACAGGTAAGAGACTAGAGACCTAGAGCGTAGAGAAACAGAGCTGAAGATGAC 2214  
Qy 2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTCAGAGCTCGAAG 2172  
Db 2215 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTCAGAGCTCGAAG 2274  
Qy 2173 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGCACTGAGAAACAGAAAGAG 2232  
Db 2275 AAACATCGGAGGCCCAAGTCTCAGCCAGCAGCTAGAGTGCACTGAGAAACAGAAAGAG 2334  
Qy 2233 CAGCACTATGAGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGCCTGCTGAC 2292  
Db 2335 CAGCACTATGAGAAAGATTAAGTGTGGCAATCAGATAAAGAAAGCCTGCTGAC 2394  
Qy 2293 AAGGACACTGGAGACATGATGCGAGACACGAGAGGAGGCCCATGAGAGGCGCAA 2352  
Db 2395 AAGGACACTGGAGACATGATGCGAGACACGAGAGGAGGCCCATGAGAGGCGCAA 2454  
Qy 2353 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGATTCGAATCCAGATCAGATCCCTGGAA 2412  
Db 2455 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGATTCGAATCCAGATCAGATCCCTGGAA 2514  
Qy 2413 CAGAGATTTGTGAACTGTCTGAAGCGCAATAAATCTTGACGCAAAATGAGAGTCTTTTACC 2472  
Db 2515 CAGAGATTTGTGAACTGTCTGAAGCGCAATAAATCTTGACGCAAAATGAGAGTCTTTTACC 2574  
Qy 2473 CAAAGGAAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532  
Db 2575 CAAAGGAAACATGAAGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2634  
Qy 2533 CTGGAGACACAGGCTGGAGTTGGAGGCCAGAACCGGAAACTGGAGAGCAGCTGGAG 2592  
Db 2635 CTGGAGACACAGGCTGGAGTTGGAGGCCAGAACCGGAAACTGGAGAGCAGCTGGAG 2694

Qy 2593 AAGATCAGCCACCAAGACCACTGATCAAGAAATCGGCTGCTGGAATCGGAGACAAAGATTG 2652  
Db 2695 AAGATCAGCCACCAAGACCACTGATCAAGAAATCGGCTGCTGGAATCGGAGACAAAGATTG 2754  
Qy 2653 CGGAGGTCTAGTCTTAGAGACGAGGAGCGAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2712  
Db 2755 CGGAGGTCTAGTCTTAGAGACGAGGAGCGAAACTGGAGCTCAAGCGCCAGCTCACAGAG 2814  
Qy 2713 CTACAGCTCTCCCTGCAAGGAGCGGAGTCAAGTTGACAGCCCTGCAAGGCTGCAAGGCG 2772  
Db 2815 CTACAGCTCTCCCTGCAAGGAGCGGAGTCAAGTTGACAGCCCTGCAAGGCTGCAAGGCG 2874  
Qy 2773 GCCCTGAGAGCGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCAAGCT 2832  
Db 2875 GCCCTGAGAGCGCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACCAAGCT 2934  
Qy 2833 GAAGAGAGATTCAGGCACTCACGGCACATAGAGATGAATAATCCAGCGCAAAATTTGATCT 2892  
Db 2935 GAAGAGAGATTCAGGCACTCACGGCACATAGAGATGAATAATCCAGCGCAAAATTTGATCT 2994  
Qy 2893 CTTTCGTAAACAGCTGTACTGTAATCAGAGCTCGAGGAGCAGCTTAAACAGCTGACCGAG 2952  
Db 2995 CTTTCGTAAACAGCTGTACTGTAATCAGAGCTCGAGGAGCAGCTTAAACAGCTGACCGAG 3054  
Qy 2953 GACAAACGCTGAACCTCAAACCAAACTTCTACTTTGTCAAAACAACCTCGATGAGGCTTCT 3012  
Db 3055 GACAAACGCTGAACCTCAAACCAAACTTCTACTTTGTCAAAACAACCTCGATGAGGCTTCT 3114  
Qy 3013 GGCGCCCAACAGCAGATTTGTAACACTCGGAAGTGAAGTGACCATCTCCGCCGGGAGATC 3072  
Db 3115 GGCGCCCAACAGCAGATTTGTAACACTCGGAAGTGAAGTGACCATCTCCGCCGGGAGATC 3174  
Qy 3073 ACGGAAACGAGAGATGCAAGCTTACAGCCAGAAAGCAACAGATGGAGGCTCTGAAGACCA 3132  
Db 3175 ACGGAAACGAGAGATGCAAGCTTACAGCCAGAAAGCAACAGATGGAGGCTCTGAAGACCA 3234  
Qy 3133 TGCAACCATGCTGAGAGAAACAGGTCATGGATTGGAGGCCCTTAAACGATGAGCTGTAGAA 3192  
Db 3235 TGCAACCATGCTGAGAGAAACAGGTCATGGATTGGAGGCCCTTAAACGATGAGCTGTAGAA 3294  
Qy 3193 AAAGAGCGGAGTGGGAGGCTGAGGAGCGTCTCGGCTGATGAGAAATCCCAAGTTTGAG 3252  
Db 3295 AAAGAGCGGAGTGGGAGGCTGAGGAGCGTCTCGGCTGATGAGAAATCCCAAGTTTGAG 3354  
Qy 3253 TGTCCGGTTCCAGAGCTGCAAGAAATCTCGAACACCGAGAAACAGAGCAGGCGAGAGCC 3312  
Db 3355 TGTCCGGTTCCAGAGCTGCAAGAAATCTCGAACACCGAGAAACAGAGCAGGCGAGAGCC 3414  
Qy 3313 GATCAGCGGATCACCGAGTCTCCCGAGGCTGGAGCTGGAGTGAAGGAGCAACAGGCT 3372  
Db 3415 GATCAGCGGATCACCGAGTCTCCCGAGGCTGGAGCTGGAGTGAAGGAGCAACAGGCT 3474  
Qy 3373 GAGATTTCTCGTCTGCAAGCGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3432  
Db 3475 GAGATTTCTCGTCTGCAAGCGGCTCTCAAAGAGCAGAAAGCTGAAGGCCGAGAGCTCTCT 3534  
Qy 3433 GACAAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAATGCCGAGCTTA 3492  
Db 3535 GACAAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAATGCCGAGCTTA 3594  
Qy 3493 CAGCAGAGCTGGAGACTGAACAGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGCCAAA 3552  
Db 3595 CAGCAGAGCTGGAGACTGAACAGAGAGCTCAAAACAGAGGCTTCTGGAAGAGCAGCCAAA 3654  
Qy 3553 TTACAGCAGCAGATGGAACCTGCGAGAAAAATCAACATTTCCGCTCTGACTCAAGGCTGAA 3612  
Db 3655 TTACAGCAGCAGATGGAACCTGCGAGAAAAATCAACATTTCCGCTCTGACTCAAGGCTGAA 3714  
Qy 3613 GAACTCTAGATCGGCTGATCTACTGAAGCAGAAAGAGTGAATTTGGAGTATCAGCTG 3672  
Db 3715 GAACTCTAGATCGGCTGATCTACTGAAGCAGAAAGAGTGAATTTGGAGTATCAGCTG 3774

QY 3673 GAAACATTTCAGGTTCTCTATTCTCATGAAAGGTGAATGGAAGCGCACTATTTCTCAA 3732  
Db 3775 GAAACATTTCAGGTTCTCTATTCTCATGAAAGGTGAATGGAAGCGCACTATTTCTCAA 3834  
QY 3733 CAACCAAACTCATTTGATTTTCTGCAAGCCAAATGACCAACTGCTGTAAGAAAGAAAAG 3792  
Db 3835 CAACCAAACTCATTTGATTTTCTGCAAGCCAAATGACCAACTGCTGTAAGAAAGAAA-- 3892  
QY 3793 GGTTTATTTAGTCGACGGAAGAGGCCCTGCTTTTACCACACAGGTTCTCTGCGAGTAC 3852  
Db 3893 -----AGTTTCTCTGCGAGTAC 3909  
QY 3853 AATGAGCTGAAGTGGCCCTGGAGAAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3912  
Db 3910 AATGAGCTGAAGTGGCCCTGGAGAAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3969  
QY 3913 CTTCAAGAACCCGATCGAGTCCGGTCCGCCGGGAGGAAGCTGCCACCGCAAGCA 3972  
Db 3970 CTTCAAGAACCCGATCGAGTCCGGTCCGCCGGGAGGAAGCTGCCACCGCAAGCA 4029  
QY 3973 ACGGACCAACCAACCATCCACGCCAGCACCGCGAGCAGCAGATCGCCATGTCGCGC 4032  
Db 4030 ACGGACCAACCAACCATCCACGCCAGCACCGCGAGCAGCAGATCGCCATGTCGCGC 4089  
QY 4033 ATCGTGGGTCGCCAGACACCGCCAGTCGATGAGCTGCTGCGCCCGCCATCCAGC 4092  
Db 4090 ATCGTGGGTCGCCAGACACCGCCAGTCGATGAGCTGCTGCGCCCGCCATCCAGC 4149  
QY 4093 CGCAGAAAGAGTCTCACTCCAGAGGAATTTAGTCGGCGCTTTAAGGAAAGCATGAC 4152  
Db 4150 CGCAGAAAGAGTCTCACTCCAGAGGAATTTAGTCGGCGCTTTAAGGAAAGCATGAC 4209  
QY 4153 CACAATATTTCTCACCAGATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTGCTGTG 4212  
Db 4210 CACAATATTTCTCACCAGATTCAACGTAGGACTGAACATCGAGCCACAAAGTGTGCTGTG 4269  
QY 4213 TGTCTGGATACCGTGACATTTGACCCAGGACATCCAAATGCTCGAATGTGAGTGATG 4272  
Db 4270 TGTCTGGATACCGTGACATTTGACCCAGGACATCCAAATGCTCGAATGTGAGTGATG 4329  
QY 4273 TGTCAACCCAAAGTCTCCAGTCTTGCAGCACCTCGCGCTTGCCTGCTCAATATGCC 4332  
Db 4330 TGTCAACCCAAAGTCTCCAGTCTTGCAGCACCTCGCGCTTGCCTGCTCAATATGCC 4389  
QY 4333 ACACATTCACAGAGGCTTTCGCGTGACAAATGAACCTCCCAAGTCTCAGACCAAG 4392  
Db 4390 ACACATTCACAGAGGCTTTCGCGTGACAAATGAACCTCCCAAGTCTCAGACCAAG 4449  
QY 4393 GAGCCACAGCAGAGCTTGACCTGGAAGGCTGATGAAGTCCCGAGGAATTAACAAAGCA 4452  
Db 4450 GAGCCACAGCAGAGCTTGACCTGGAAGGCTGATGAAGTCCCGAGGAATTAACAAAGCA 4509  
QY 4453 GGAACAGAGGCTGGACAGGAAGTACATTTGCTGGAGGATCAAAAGTCTCTATTTAT 4512  
Db 4510 GGAACAGAGGCTGGACAGGAAGTACATTTGCTGGAGGATCAAAAGTCTCTATTTAT 4569  
QY 4513 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGCGCTTCCC 4572  
Db 4570 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAGAAATTTGAGCTGCGCTTCCC 4629  
QY 4573 GACGGGATGTATCTATTATCATGGTCCGCTGCTGCTTCCGAATCGCAATACAGCCAAA 4632  
Db 4630 GACGGGATGTATCTATTATCATGGTCCGCTGCTGCTTCCGAATCGCAATACAGCCAAA 4689  
QY 4633 GCA----- 4635  
Db 4690 GCAGATGTCCCATACATCTGAAGTGAATCTCACCGCACACACCTGCTGGCCGGG 4749  
QY 4636 ----- 4635  
Db 4750 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4809  
QY 4636 -----GAAAAAGCAGAAAGCTGATGCTAAACTG 4662

Db 4810 GAATCAGTTGTCGAGGTGGGAGAGTTTCTAGGAAAGAAAGCAAGCTGATGCTAAACTG 4869  
QY 4663 CTTGAAACTCCCTGCTGCTGAAACTGAAAGGTGATGACCGTCTTAGACATGAACTGCACGCTG 4722  
Db 4870 CTTGAAACTCCCTGCTGAAACTGAAAGGTGATGACCGTCTTAGACATGAACTGCACGCTG 4929  
QY 4723 CCGTTCACTGACAGGTGGTGTGCTGGTGGCCACGAGGAGGGCTCTAGCCCTGAATGTC 4782  
Db 4930 CCGTTCACTGACAGGTGGTGTGCTGGTGGCCACCGAGGAGGGCTCTAGCCCTGAATGTC 4989  
QY 4783 TTGAAAACTCCCTAAACCCATGTCCAGGAATTCGAGCAGTCTTCCAAATTTATATTATC 4842  
Db 4990 TTGAAAACTCCCTAAACCCATGTCCAGGAATTCGAGCAGTCTTCCAAATTTATATTATC 5049  
QY 4843 AAGGACCTGGAGAGCTTACTCATGATAGCAGGAGAAAGCGGGCACTGTGCTTGTGGAC 4902  
Db 5050 AAGGACCTGGAGAGCTTACTCATGATAGCAGGAGAAAGCGGGCACTGTGCTTGTGGAC 5109  
QY 4903 GTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCCAGCCGACATCTCA 4962  
Db 5110 GTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCCAGCCGACATCTCA 5169  
QY 4963 CCCAACATTTTGAAGCTGTCAAGGGTGCACATTTGTTGGGCGAGCAAGATTGAAC 5022  
Db 5170 CCCAACATTTTGAAGCTGTCAAGGGTGCACATTTGTTGGGCGAGCAAGATTGAAC 5229  
QY 5023 GGGCTCTGCATCTGTGAGCAGCATGCCAGCAAAAGTGTGCTTCTCGCTCAACAGAAAC 5082  
Db 5230 GGGCTCTGCATCTGTGAGCAGCATGCCAGCAAAAGTGTGCTTCTCGCTCAACAGAAAC 5289  
QY 5083 CTCAGCAAAATCTGCATCCGGAAGAGATAGACCTTCAGAGCCCTGACGTGTATCCAC 5142  
Db 5290 CTCAGCAAAATCTGCATCCGGAAGAGATAGACCTTCAGAGCCCTGACGTGTATCCAC 5349  
QY 5143 TTCAACCAATTCAGTATCTCTATTGGNACCAATAAATTTACGAAATCGACATGAGCAG 5202  
Db 5350 TTCAACCAATTCAGTATCTCTCAATTGGAACCAATAAATTTACGAAATCGACATGAGCAG 5409  
QY 5203 TACACGCTCCAGGAATTCCTGGATAGAATGACATTCCTTTGGCAGCTGTGTGTTGCC 5262  
Db 5410 TACACGCTCCAGGAATTCCTGGATAGAATGACATTCCTTTGGCAGCTGTGTGTTGCC 5469  
QY 5263 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGAGGGCAGCAGAG 5322  
Db 5470 GCCTCTTCCAAAGCTTCCCTGTCTCAATCGTGCAGGTGAACAGCGAGGGCAGCAGAG 5529  
QY 5323 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC 5382  
Db 5530 GAGTACTTGTGTGTTTCCACGAATTTGGAGTGTTCGTGGATTTCTTACGGAAGACGTAGC 5589  
QY 5383 CGCACAGACGATCTCAAGTGGAGTGCCTTACCTTTTGGCCCTTGCCTTACAGAGAACCTAT 5442  
Db 5590 CGCACAGACGATCTCAAGTGGAGTGCCTTACCTTTTGGCCCTTGCCTTACAGAGAACCTAT 5649  
QY 5443 CTGTTTGTGACCCACATTCACCTCAGTTCAGTAAATTTAGATCCAGGACGCTCTCAGCA 5502  
Db 5650 CTGTTTGTGACCCACATTCACCTCAGTTCAGTAAATTTAGATCCAGGACGCTCTCAGCA 5709  
QY 5503 GGGACCCCTCCCGAGCGTACCTGGACATCCCGAACCCGCGCTTACCTGGGCGCTTGCAT 5562  
Db 5710 GGGACCCCTCCCGAGCGTACCTGGACATCCCGAACCCGCGCTTACCTGGGCGCTTGCAT 5769  
QY 5563 TCCCTCAGGACGATTTTACTTGGCGTCTCTATACAGGATAAATTAAGGTGCTTCTGCTC 5622  
Db 5770 TCCCTCAGGACGATTTTACTTGGCGTCTCTATACAGGATAAATTAAGGTGCTTCTGCTC 5829  
QY 5623 AAGGGAACCTCTGTGAAGGAGTCCGGCACTGAAACACACCGGGGCGCTTCCACCTCCGC 5682  
Db 5830 AAGGGAACCTCTGTGAAGGAGTCCGGCACTGAAACACACCGGGGCGCTTCCACCTCCGC 5889  
QY 5683 AGCAGCCCAACAGGAGGAGCCCAACCGTACACAGCAGCATACCAAGCGCGTGGCC 5742

```
Db 5890 AGCAGCCCAACAAGCGAGGCCACCACGCTACAAGAGCACATCACCAAGCGGTGGCC 5949
Qy 5743 TCCAGCCAGCGCCGCCGAAGCGCCCGACCGACCGCGAGAGCCAGCACACCCACCGC 5802
Db 5950 TCCAGCCAGCGCCGCCGAAGCGCCCGACCGACCGCGAGAGCCAGCACACCCACCGC 6009
Qy 5803 TACCGCGAGGGCGGACCGAGCTGGCGAGGGAACAAGTCTCTGGCGCGCCCTGGAGCGA 5862
Db 6010 TACCGCGAGGGCGGACCGAGCTGGCGAGGGAACAAGTCTCTGGCGCGCCCTGGAGCGA 6069
Qy 5863 GAGAGTCTCCCGCGCGGATGCTCAGCAGCGCGGAGAGCGGTCCCGCGGAGGCTGT 5922
Db 6070 GAGAGTCTCCCGCGCGGATGCTCAGCAGCGCGGAGAGCGGTCCCGCGGAGGCTGT 6129
Qy 5923 GAAGCAGCAGCAGCGCGCGGTGCTGCTGGCGGAGCGGTGAGCAGCCCGGTCTCCAGGTG 5982
Db 6130 GAAGCAGCAGCAGCGCGCGGTGCTGCTGGCGGAGCGGTGAGCAGCCCGGTCTCCAGGTG 6189
Qy 5983 AACAGGGAGGCGAGAGTGCCTCTCAAGTTTCAAGGTTTAACTGTCACTTATTAT 6042
Db 6190 AACAGGTTCTGGACCAAGTCTTCAGTATAAATCTCAGCCAGAGAAAAACCACTCTCATCT 6249

RESULT 5
AAL55214
ID AAL55214 standard; DNA; 6165 BP.
XX
AC AAL55214;
XX
DT 01-VAY-2003 (first entry)
XX
DE Human CR1K encoding DNA sequence, SEQ ID No 1.
XX
KW Anorectic; hypotensive; cardiant; antilipaeamic; cerebroprotective;
KW antitout; osteopathic; antiarthritic; cytostatic; antidepressant;
KW immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;
KW neuroprotective; antiinflammatory; antidiabetic; analgesic;
KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;
KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;
KW central nervous system disorder; chronic obstructive pulmonary disease;
KW diabetes; pain; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..6165
FT FT /*tag= a
FT FT /product= "Human CR1K protein"
XX
PN WO2003004523-A1.
XX
PD 16-JAN-2003.
XX
PF 28-JUN-2002; 2002WO-EP007156.
XX
PR 02-JUL-2001; 2001US-0301841P.
XX
PR 11-DEC-2001; 2001US-033851P.
XX
PR 25-APR-2002; 2002US-0375014P.
XX
PA {FARB } BAYER AG.
XX
PI Zhu Z;
XX
WI WPI; 2003-221576/21.
XX
DR P-PSDB; AAO26959.
XX
XX
XX New human citron rho/rac-interacting kinase (CR1K) polypeptide and
XX polynucleotide, useful in preventing, ameliorating or treating diseases
XX associated with human CR1K dysfunction, e.g. obesity, diabetes or
XX Alzheimer's disease.
XX
XX Example 1; Fig 1; 237pp; English.
XX
XX
```

```
CC The invention relates to an isolated polynucleotide encoding a human
CC citron rho/rac-interacting kinase polypeptide. The isolated
CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the
CC specification. The human citron rho/rac-interacting kinase (CR1K)
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or
CC treating diseases associated with human CR1K dysfunction such as obesity
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery
CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of
CC cancer including endometrial, breast, prostate and colon cancer),
CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's
CC disease), chronic obstructive pulmonary disease, or diabetes. These can
CC also be used to treat pain associated with the disorders. The human CR1K
CC polypeptide is also useful in diagnostic assays or in genetic testing.
CC The expression vector or the reagent is useful in preparing a medicament
CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a
CC central nervous system disorder, or chronic obstructive pulmonary
CC disease. The fusion protein is useful for generating antibodies against a
CC CR1K polypeptide and for use in various assay systems. The methods are
CC useful in producing and detecting the polynucleotide and polypeptide and
CC in screening for agents that modulate the activity of the human CR1K
CC polypeptide. This polynucleotide sequence represents a DNA sequence
CC encoding a human CR1K protein of the invention
XX
```

Sequence 6165 BP; 1735 A; 1549 C; 1680 G; 1201 T; 0 U; 0 Other;

Query Match 92.0%; Score 5666.2; DB 8; Length 6165;  
Best Local Similarity 95.9%; Pred. No. 0;  
Matches 5947; Conservative 0; Mismatches 13; Indels 243; Gaps 3;

```
Qy 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATCGTGGTGTGACCCCAT 60
Db 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCCTTTGGATCGTGGTGTGACCCCAT 60
Qy 61 GCCAGCGGGCGCTCCAGGCTGAATCTCTTCCAGGGGAAACACACCTTTATGACTCAA 120
Db 61 GCCAGCGGGCGCTCCAGGCTGAATCTCTTCCAGGGGAAACACACCTTTATGACTCAA 120
Qy 121 CAGCAGATGCTCTCTTTCCGAGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
Db 121 CAGCAGATGCTCTCTTTCCGAGAGAGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
Qy 181 GAATGCAGTCAGCTGCTCTGATGAAGATTAAGCAGCTGAGCACTTTGTCGGGAAGTAT 240
Db 181 GAATGCAGTCAGCTGCTCTGATGAAGATTAAGCAGCTGAGCACTTTGTCGGGAAGTAT 240
Qy 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGAGACTTGAAGTCAGA 300
Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGGCAAGAGACTTGAAGTCAGA 300
Qy 301 AGTCTTGTAGTTGTGGTCACTTTGCTGAAGTCAGCTGAGTGAAGAGAAAGCAACCGGG 360
Db 301 AGTCTTGTAGTTGTGGTCACTTTGCTGAAGTCAGCTGAGTGAAGAGAAAGCAACCGGG 360
Qy 361 GACATCTATGCTATGAAGTGAAGAGAGAGAGGCTTTATTGGCCAGGAGCAGGTTTCA 420
Db 361 GACATCTATGCTATGAAGTGAAGAGAGAGAGGCTTTATTGGCCAGGAGCAGGTTTCA 420
Qy 421 TTTTTCAGGAGAGCGGAGACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480
Db 421 TTTTTCAGGAGAGCGGAGACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480
Qy 481 CAGTATGCCCTTCAGGACAAAATCACCCTTTATCTGATGGAGGAATATCAGCTGGAGGG 540
Db 481 CAGTATGCCCTTCAGGACAAAATCACCCTTTATCTGATGGAGGAATATCAGCTGGAGGG 540
Qy 541 GACTTGTCTGCTTTTGAATAGATATGAGGACAGTATAGATGAAAACCTGTATACAGTTT 600
Db 541 GACTTGTCTGCTTTTGAATAGATATGAGGACAGTATAGATGAAAACCTGTATACAGTTT 600
Qy 601 TACCTAGCTGAGCTGATTTGGCTGTTTCAGAGGCTTCATCTGATGGGATACGTCATCGA 660
Db 601 TACCTAGCTGAGCTGATTTGGCTGTTTCAGAGGCTTCATCTGATGGGATACGTCATCGA 660
```

QY 661 GACATCAAGCCTGAGAACATCTCGTTGACCGCACAGGACACATCAAGCTGTGGATTTT 720  
DB 661 GACATCAAGCCTGAGAACATCTCGTTGACCGCACAGGACACATCAAGCTGTGGATTTT 720  
QY 721 GGATCTCCCGGAAATGAATTCACACAGATGGTGAATGCCAACTCCGATTTGGACC 780  
DB 721 GGATCTCCCGGAAATGAATTCACACAGATGGTGAATGCCAACTCCGATTTGGACC 780  
QY 781 CCAGATTACATGGCTCCTGAAAGTCTGACTGTGATGAACCGGGATGGAAGGCACTTAC 840  
DB 781 CCAGATTACATGGCTCCTGAAAGTCTGACTGTGATGAACCGGGATGGAAGGCACTTAC 840  
QY 841 GGCTGACATGTGACTGGTGGTCACTGAGTGGGCTGATTCCTATGAGATGATTTATGGGAGA 900  
DB 841 GGCTGACATGTGACTGGTGGTCACTGAGTGGGCTGATTCCTATGAGATGATTTATGGGAGA 900  
QY 901 TCCCTCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAACAATTAATGAATTTCAAGCGG 960  
DB 901 TCCCTCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAACAATTAATGAATTTCAAGCGG 960  
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCGAATCAAAGC 1020  
DB 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATTCGAATCAAAGC 1020  
QY 1021 TTGTTGCGCGCCAGAAAGAGACTGAAGTTTGAAGGCTTTTGCTGCCATCCCTTTCTTC 1080  
DB 1021 TTGTTGCGCGCCAGAAAGAGACTGAAGTTTGAAGGCTTTTGCTGCCATCCCTTTCTTC 1080  
QY 1081 TCTAAAATTTGACTGGAAACAATTCGTAACTCTCTCCCTCCCTCTGTTCCACCTCAAG 1140  
DB 1081 TCTAAAATTTGACTGGAAACAATTCGTAACTCTCTCTCCCTCTGTTCCACCTCAAG 1140  
QY 1141 TCTGACATGACACTTCCAAATTTTGATGAACACAGAGAAGATTCGTGGTTCATCTCT 1200  
DB 1141 TCTGACATGACACTTCCAAATTTTGATGAACACAGAGAAGATTCGTGGTTCATCTCT 1200  
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGGGGTTTCG 1260  
DB 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCGGTTTGGGGTTTCG 1260  
QY 1261 TACAGCAAGCCTGGGATCTTGTGTAGATCTGATCTGTTGTGTGGGTTCGACTCC 1320  
DB 1261 TACAGCAAGCCTGGGATCTTGTGTAGATCTGATCTGTTGTGTGGGTTCGACTCC 1320  
QY 1321 CCGTGCAGACTAGCTCCATGGAAGAAACTTCTCATCAAAGCAAGAGCTACAAGAC 1380  
DB 1321 CCGTGCAGACTAGCTCCATGGAAGAAACTTCTCATCAAAGCAAGAGCTACAAGAC 1380  
QY 1381 TCTCAGGACAAGTGTCAAGAATGGAGAGGAAATGACCCGTTTACATCGGAGTGTCA 1440  
DB 1381 TCTCAGGACAAGTGTCAAGAATGGAGAGGAAATGACCCGTTTACATCGGAGTGTCA 1440  
QY 1441 GAGGTGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTCAAGGCTCTGAGACTCAGAGA 1500  
DB 1441 GAGGTGAGGCTGTGCTTAGTCAGAGGAGGTGGAGCTCAAGGCTCTGAGACTCAGAGA 1500  
QY 1501 TCGCTCTGGAGCAGGACCTTGCTACCTACATCACAGAAATGCAAGTGTAAAGCGAAGT 1560  
DB 1501 TCGCTCTGGAGCAGGACCTTGCTACCTACATCACAGAAATGCAAGTGTAAAGCGAAGT 1560  
QY 1561 TTGAGCAAGCAGGATGAGGTGTCCAGGAGGATGACAAAGCACTGAGCTTCCAT 1620  
DB 1561 TTGAGCAAGCAGGATGAGGTGTCCAGGAGGATGACAAAGCACTGAGCTTCCAT 1620  
QY 1621 GATATCAGAGAGCAGAGCCGAGCTTCCAAAGAAATCAAAGACAGGAGTACCAGGCTCAA 1680  
DB 1621 GATATCAGAGAGCAGAGCCGAGCTTCCAAAGAAATCAAAGACAGGAGTACCAGGCTCAA 1680  
QY 1681 GTGGAAGAAATGAGGTGTGATGAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740  
DB 1681 GTGGAAGAAATGAGGTGTGATGAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740

QY 1741 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAATTCAG 1800  
DB 1741 CGGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAATTCAG 1800  
QY 1801 CGGAAGCGCAGAGATGTACGATAAACTGTTGAAGCTTAAGGATCAAGGAGACCTGAA 1860  
DB 1801 CGGAAGCGCAGAGATGTACGATAAACTGTTGAAGCTTAAGGATCAAGGAGACCTGAA 1860  
QY 1861 GTGGAGAAATATCGAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
DB 1861 GTGGAGAAATATCGAAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAG 1920  
QY 1921 CTCCAAGAGAACTGGAGAGGCTT----- 1944  
DB 1921 CTCCAAGAGAACTGGAGAGGCTTAAAGCCAGCACGAGGCCACCGAGCTGCTCAG 1980  
QY 1945 -----GCAAGGAGCGAGCCGAGAGGAGCTGGAGAGCTGCGAAGCCGAGAG 1992  
DB 1945 -----GCAAGGAGCGAGCCGAGAGGAGCTGGAGAGCTGCGAAGCCGAGAG 1992  
QY 1993 AATATCCCGCAGGCAAGGAGCGAGCCGAGAGGAGCTGGAGAGCTGCGAAGCCGAGAG 2040  
DB 1993 AATATCCCGCAGGCAAGGAGCGAGCCGAGAGGAGCTGGAGAGCTGCGAAGCCGAGAG 2040  
QY 2052 GATCTCTTGAAGCATCAGAAAGAGCTGGTGAAGCTGAGGAAACGCCGCCATTTCTCTG 2052  
DB 2052 GATCTCTTGAAGCATCAGAAAGAGCTGGTGAAGCTGAGGAAACGCCGCCATTTCTCTG 2052  
QY 2100 GATCTCTTGAAGCATCAGAAAGAGCTGGTGAAGCTGAGGAAACGCCGCCATTTCTCTG 2100  
DB 2100 GATCTCTTGAAGCATCAGAAAGAGCTGGTGAAGCTGAGGAAACGCCGCCATTTCTCTG 2100  
QY 2112 GAGAACCAAGGTAAAGAGACTAGAGACCATGAGCGTAGAGAAACAGACTGAAGGATGAC 2112  
DB 2112 GAGAACCAAGGTAAAGAGACTAGAGACCATGAGCGTAGAGAAACAGACTGAAGGATGAC 2112  
QY 2160 GAGAACCAAGGTAAAGAGACTAGAGACCATGAGCGTAGAGAAACAGACTGAAGGATGAC 2160  
DB 2160 GAGAACCAAGGTAAAGAGACTAGAGACCATGAGCGTAGAGAAACAGACTGAAGGATGAC 2160  
QY 2172 ATCCAGCAAAATCCCAACAGATCCAGCAGATGCTGATAAAAATTTCTGGAGCTCGAAGAG 2172  
DB 2172 ATCCAGCAAAATCCCAACAGATCCAGCAGATGCTGATAAAAATTTCTGGAGCTCGAAGAG 2172  
QY 2232 AAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTAGAGTGCACCTGTAACAGAGAGAG 2232  
DB 2232 AAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTAGAGTGCACCTGTAACAGAGAGAG 2232  
QY 2280 AAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTAGAGTGCACCTGTAACAGAGAGAG 2280  
DB 2280 AAACATCGGAGGCGCCAAAGTCTCAGCCAGCAGCTAGAGTGCACCTGTAACAGAGAGAG 2280  
QY 2292 CAGCACTATGAGGAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2292  
DB 2292 CAGCACTATGAGGAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2292  
QY 2340 CAGCACTATGAGGAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2340  
DB 2340 CAGCACTATGAGGAAAAGATTAAAGTGTGGAACAATCAGATAAAGAAAGACCTGGCTGAC 2340  
QY 2352 AAGGAGACACTGGAGAACATGATGACAGACACAGAGGAGGAGGCCATGAGAGGCGAAA 2352  
DB 2352 AAGGAGACACTGGAGAACATGATGACAGACACAGAGGAGGAGGCCATGAGAGGCGAAA 2352  
QY 2400 AAGGAGACACTGGAGAACATGATGACAGACACAGAGGAGGAGGCCATGAGAGGCGAAA 2400  
DB 2400 AAGGAGACACTGGAGAACATGATGACAGACACAGAGGAGGAGGCCATGAGAGGCGAAA 2400  
QY 2412 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGATTCAGATCCAGATCCCTGGAA 2412  
DB 2412 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGATTCAGATCCAGATCCCTGGAA 2412  
QY 2460 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTTGACAGCAATAGCAGTCTTTTACC 2460  
DB 2460 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTTGACAGCAATAGCAGTCTTTTACC 2460  
QY 2472 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTTGACAGCAATAGCAGTCTTTTACC 2472  
DB 2472 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTTGACAGCAATAGCAGTCTTTTACC 2472  
QY 2520 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTTGACAGCAATAGCAGTCTTTTACC 2520  
DB 2520 CAGAGGATTTGGAACTGTCTGAAGCCAAATAAATTTGACAGCAATAGCAGTCTTTTACC 2520  
QY 2532 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532  
DB 2532 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2532  
QY 2580 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580  
DB 2580 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580  
QY 2592 CTGAGACACAGGCTGGAAAGTTGGAGGCCGAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2592  
DB 2592 CTGAGACACAGGCTGGAAAGTTGGAGGCCGAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2592  
QY 2640 CTGAGACACAGGCTGGAAAGTTGGAGGCCGAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2640  
DB 2640 CTGAGACACAGGCTGGAAAGTTGGAGGCCGAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2640  
QY 2652 AAGATCAGCCCAACAGACACATGATGAGAAATCGGCTGCTGAACTGGAGACAGATTTG 2652  
DB 2652 AAGATCAGCCCAACAGACACATGATGAGAAATCGGCTGCTGAACTGGAGACAGATTTG 2652  
QY 2700 AAGATCAGCCCAACAGACACATGATGAGAAATCGGCTGCTGAACTGGAGACAGATTTG 2700  
DB 2700 AAGATCAGCCCAACAGACACATGATGAGAAATCGGCTGCTGAACTGGAGACAGATTTG 2700  
QY 2712 CGGAGGCTCAGTCTAGAGCACGAGGAGCAGAAACTGAGGCTCAAGGCCAGCTCACAGAG 2712  
DB 2712 CGGAGGCTCAGTCTAGAGCACGAGGAGCAGAAACTGAGGCTCAAGGCCAGCTCACAGAG 2712  
QY 2760 CGGAGGCTCAGTCTAGAGCACGAGGAGCAGAAACTGAGGCTCAAGGCCAGCTCACAGAG 2760  
DB 2760 CGGAGGCTCAGTCTAGAGCACGAGGAGCAGAAACTGAGGCTCAAGGCCAGCTCACAGAG 2760  
QY 2772 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTGACAGCTGACAGCTGACCGGGG 2772  
DB 2772 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTGACAGCTGACAGCTGACCGGGG 2772  
QY 2820 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTGACAGCTGACAGCTGACCGGGG 2820  
DB 2820 CTACAGCTCTCCCTGACGAGCGGAGTCAAGTTGACAGCTGACAGCTGACCGGGG 2820  
QY 2832 GCGCTGAGAGCCAGCTTTCGCCAGGCGAAGACAGAGCTGGAAAGAGACCACAGCAAGCT 2832  
DB 2832 GCGCTGAGAGCCAGCTTTCGCCAGGCGAAGACAGAGCTGGAAAGAGACCACAGCAAGCT 2832

Db ||||| 2821 GGCCTGGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGCAAGCT 2880  
QY ||||| 2833 GAAGGAGATCCAGCACTCAGGCACATAGAGATGAATCCAGCGCAATTTTCATGCT 2892  
Db ||||| 2881 GAAGGAGATCCAGGCACTCAGGCACATAGAGATGAATCCAGCGCAATTTTCATGCT 2940  
QY ||||| 2893 CTTTCGTAAACAGCTGTACTGTAAATCAGACCTCGGAGGAGCAGCTAAACAGCTGACCGAG 2952  
Db ||||| 2941 CTTTCGTAAACAGCTGTACTGTAAATCAGACCTCGGAGGAGCAGCTAAACAGCTGACCGAG 3000  
QY ||||| 2953 GACACGCTGAATCMAACCAACAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 3012  
Db ||||| 3001 GACACGCTGAATCMAACCAACAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 3060  
QY ||||| 3013 GCGGCCAACGAGAGATTCTACAACTGCGAAGTGAAGTGGACATCTCCGCCGGGAGATC 3072  
Db ||||| 3061 GCGGCCAACGAGAGATTCTACAACTGCGAAGTGAAGTGGACATCTCCGCCGGGAGATC 3120  
QY ||||| 3073 ACGGAACGAGATGACGCTTACAGCCGAGAGCAAAACGATGGAGGCTCTGGAAGCCACG 3132  
Db ||||| 3121 ACGGAACGAGATGACGCTTACAGCCGAGAGCAAAACGATGGAGGCTCTGGAAGCCACG 3180  
QY ||||| 3133 TGCACCATCTCGAGGAAACAGGTCTATGGATTGAGAGGCCCTAAACGATGAGCTGCTAGAA 3192  
Db ||||| 3181 TGCACCATCTCGAGGAAACAGGTCTATGGATTGAGAGGCCCTAAACGATGAGCTGCTAGAA 3240  
QY ||||| 3193 AAAGAGCGCAGTGGAGGCGCTGGAGGCGCTCTGGGATGAGAAATCCAGTTTGA 3252  
Db ||||| 3241 AAAGAGCGCAGTGGAGGCGCTGGAGGCGCTCTGGGATGAGAAATCCAGTTTGA 3300  
QY ||||| 3253 TGTCCGGTTTCGAGAGCTGAGAGATGCTGGACACCGAGAAACAGAGCGCGAGGCC 3312  
Db ||||| 3301 TGTCCGGTTTCGAGAGCTGAGAGATGCTGGACACCGAGAAACAGAGCGCGAGGCC 3360  
QY ||||| 3313 GATCAGCGGATCACCGAGTCTCGCCAGTGGTGGAGCTGGCAGTGAAGAGCACAAGGCT 3372  
Db ||||| 3361 GATCAGCGGATCACCGAGTCTCGCCAGTGGTGGAGCTGGCAGTGAAGAGCACAAGGCT 3420  
QY ||||| 3373 GAGATTCTCGCTTCGAGAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCT 3432  
Db ||||| 3421 GAGATTCTCGCTTCGAGAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGCCTCTCT 3480  
QY ||||| 3433 GACAACTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3492  
Db ||||| 3481 GACAACTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3540  
QY ||||| 3493 CAGCAGAAGCTGGAGACTGAACGAGAGCTCAACACAGAGGCTTCTGGAAGAGCAAGCCAAA 3552  
Db ||||| 3541 CAGCAGAAGCTGGAGACTGAACGAGAGCTCAACACAGAGGCTTCTGGAAGAGCAAGCCAAA 3600  
QY ||||| 3553 TTACAGCAGCATGACCTGCAGAAATACATTTCCGTCGACTCAAGGACTGGAA 3612  
Db ||||| 3601 TTACAGCAGCATGACCTGCAGAAATACATTTCCGTCGACTCAAGGACTGGAA 3660  
QY ||||| 3613 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGAGTATCAGCTG 3672  
Db ||||| 3661 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGAGTATCAGCTG 3720  
QY ||||| 3673 GAAAAATTTCAGTTCTTCTATCTCATGAAAGGTGAATGGAAGGCACTATTCTCAA 3732  
Db ||||| 3721 GAAAAATTTCAGTTCTTCTATCTCATGAAAGGTGAATGGAAGGCACTATTCTCAA 3780  
QY ||||| 3733 CAAACCAAACTCATTTGATTTCTGCAAGCAAAATGGACCAACCTGCTAAAAAGAAAAAG 3792  
Db ||||| 3781 CAAACCAAACTCATTTGATTTCTGCAAGCAAAATGGACCAACCTGCTAAAAAGAAAA-- 3838  
QY ||||| 3793 GGTTTATTTAGTCGACGGAAGAGGACCTGCTTTACCACACAGAGTTCCTCTGAGTAC 3852  
Db ||||| 3839 -----AGTTTCTCTGAGTAC 3855  
QY ||||| 3853 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAGCC 3912

Db ||||| 3856 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3915  
QY ||||| 3913 CTTTCAGAAAGACCCGCACTCGAGCTCCGCTCCGCCGGGAGGAAGCTGCCCAACGAAAGCA 3972  
Db ||||| 3916 CTTTCAGAAAGACCCGCACTCGAGCTCCGCTCCGCCGGGAGGAAGCTGCCCAACGAAAGCA 3975  
QY ||||| 3973 ACGAACCAACCAACACCCATCCAGCCAGCACCCGCGAGGAGCAGAGATCGGCATGTCGCGC 4032  
Db ||||| 3976 ACGAACCAACCAACACCCATCCAGCCAGCACCCGCGAGGAGCAGAGATCGGCATGTCGCGC 4035  
QY ||||| 4033 ATCGTGGGTCGCCAGAGCACCCAGCCAGTGCATGAGCTGTGGCCCGCCGCAATCCAGC 4092  
Db ||||| 4036 ATCGTGGGTCGCCAGAGCACCCAGCCAGTGCATGAGCTGTGGCCCGCCGCAATCCAGC 4095  
QY ||||| 4093 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGCAATGCAC 4152  
Db ||||| 4096 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGTCTTAAGGAACGCAATGCAC 4155  
QY ||||| 4153 CACAAATATCTCACCGATTCMAAGTAGGACTGAAATGCGAGCCAGGCAAAAGTGTGCTGTG 4212  
Db ||||| 4156 CACAAATATCTCACCGATTCMAAGTAGGACTGAAATGCGAGCCAGGCAAAAGTGTGCTGTG 4215  
QY ||||| 4213 TGTCTGGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCCAGTGTATG 4272  
Db ||||| 4216 TGTCTGGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCCAGTGTATG 4275  
QY ||||| 4273 TGTCAACCCCAAGTGTCCAGTGTCCAGGCCACCTGCGGCTTGCCTGCTGAATATGCC 4332  
Db ||||| 4276 TGTCAACCCCAAGTGTCCAGTGTCCAGGCCACCTGCGGCTTGCCTGCTGAATATGCC 4335  
QY ||||| 4333 ACACATTTACCGAGGCTTCTGCCGTGACAAATGAATCACTCCCGAGGTCTCCAGACCAAG 4392  
Db ||||| 4336 ACACATTTACCGAGGCTTCTGCCGTGACAAATGAATCACTCCCGAGGTCTCCAGACCAAG 4395  
QY ||||| 4393 GAGCCAGCAGCAGCTTTCGACCTGGAAGGTTGATGAAGTGCAGGAATAACAAACGA 4452  
Db ||||| 4396 GAGCCAGCAGCAGCTTTCGACCTGGAAGGTTGATGAAGTGCAGGAATAACAAACGA 4455  
QY ||||| 4453 GGACAGCAAGGCTGGGACAGAGAGTACATTTGCTCGAGGGATCAAAAGTCCATTTAT 4512  
Db ||||| 4456 GGACAGCAAGGCTGGGACAGAGAGTACATTTGCTCGAGGGATCAAAAGTCCATTTAT 4515  
QY ||||| 4513 GACAAATGAAGCAGAGAGGCTGGACAGGCGCGTGGAGAAATTTGAGCTGTGCTTCCC 4572  
Db ||||| 4516 GACAAATGAAGCAGAGAGGCTGGACAGGCGCGTGGAGAAATTTGAGCTGTGCTTCCC 4575  
QY ||||| 4573 GACGGGATGTATCTATTTATGCTGCGCTTGGTGTCTTCGAACTCGCAATACAGCCAAA 4632  
Db ||||| 4576 GACGGGATGTATCTATTTATGCTGCGCTTGGTGTCTTCGAACTCGCAATACAGCCAAA 4635  
QY ||||| 4633 GCA----- 4635  
Db ||||| 4636 GCAGATGTCCTCATACATACTGAAGATGGAATCTCACCCGACACACCCTGCTGCGCCGGG 4695  
QY ||||| 4636 ----- 4636  
Db ||||| 4696 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAACAGCGGCTGGGTACCGGCTTA 4755  
QY ||||| 4636 -----GAAAAAGCAGAAGCTGATGCTAAACTG 4662  
Db ||||| 4756 GAATCATGTTGTCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAAGCTGATGCTAAACTG 4815  
QY ||||| 4663 CTTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAAGTCAAGCTG 4722  
Db ||||| 4816 CTTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCGTCTAGACATGAAGTCAAGCTG 4875  
QY ||||| 4723 CCCTTCAGTGACCAAGTGTGTTGGTGGGACCGAGGAAGGCTCTACGCCCTGAATGTC 4782  
Db ||||| 4876 CCCTTCAGTGACCAAGTGTGTTGGTGGGACCGAGGAAGGCTCTACGCCCTGAATGTC 4935  
QY ||||| 4783 TTGAAAAACTCCCTTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTTATTTATC 4842  
Db ||||| 4936 TTGAAAAACTCCCTTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTTATTTATC 4995

QY	4843	AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTCTGTGGAC	4902	QY	5923	CAAGACAGCAGCAGGCGCGCTCCCTCGGAGCCCTGAGGACCCCGCTGTCCAGGTG	5982
DB	4996	AAGGACCTGGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACTGTCTGTGGAC	5055	DB	6076	GAAGACAGCAGCAGGCGCGCTCCCTCGGAGCCCTGAGGACCCCGCTGTCCAGGTG	6135
QY	4903	GTGAAGAAAGTGAACAGTCCCTGGCCCCAGTCCACCTGCTGCCAGCCGACATCTCA	4962	QY	5983	AACAAGGAGAGGCGAGGTGC	6005
DB	5056	GTGAAGAAAGTGAACAGTCCCTGGCCCCAGTCCACCTGCTGCCAGCCGACATCTCA	5115	DB	6136	AACAAGGTCTGGACCACTCTTC	6158
QY	4963	CCCAACATTTTGAAGCTGTCAGGCTGCCACTTGTGGGCGAGGCAAGATTGAGAC	5022	RESULT 6			
DB	5116	CCCAACATTTTGAAGCTGTCAGGCTGCCACTTGTGGGCGAGGCAAGATTGAGAC	5175	ID	ABQ78870	standard; cDNA; 6165 BP.	
QY	5023	GGGCTCTGCATCTGTGAGCAGTCCGAGCAAGTCTCATCTCTGCTACACGAAAC	5082	AC	ABQ78870;		
DB	5176	GGGCTCTGCATCTGTGAGCAGTCCGAGCAAGTCTCATCTCTGCTACACGAAAC	5235	XX	10-OCT-2002 (first entry)		
QY	5083	CTCAGCAATATCTGCATCCGGAAGAGATAGACCTCAGACCTCAGACCTGTATCCAC	5142	XX	Human kinase cDNA #1.		
DB	5236	CTCAGCAATATCTGCATCCGGAAGAGATAGACCTCAGACCTCAGACCTGTATCCAC	5295	XX	Human; kinase; enzyme; serine-threonine kinase; neotropic; cytostatic;		
QY	5143	TTCAACCAATTTACAGTATCTCTTGAACCAATTAATTTAGCAATCGACATGAAGAG	5202	KW	Citron rho-interacting kinase; gene therapy; mental disorder; cancer;		
DB	5296	TTCAACCAATTTACAGTATCTCTTGAACCAATTAATTTAGCAATCGACATGAAGAG	5355	KW	gene; ss.		
QY	5203	TACACGCTCAGGAATTCCTGGATAAGATGACATTCCTTGCCACCTGCTGTGTC	5262	OS	Homo sapiens.		
DB	5356	TACACGCTCAGGAATTCCTGGATAAGATGACATTCCTTGCCACCTGCTGTGTC	5415	FH	Key	Location/Qualifiers	
QY	5263	GCCTCTTCAACAGCTTCCTGCTCAATCTGAGGTGAACAGCGCAGGCGAGGAG	5322	FT	CDS	1. 6165	
DB	5416	GCCTCTTCAACAGCTTCCTGCTCAATCTGAGGTGAACAGCGCAGGCGAGGAG	5475	FT	variation	/*tag= a	
QY	5323	GAGTACTTGTGTGTTTCCAGCAATTTGGAGTTCCTGCGATTTTACGGAAGCTAGC	5382	FT	variation	/product= "Kinase"	
DB	5476	GAGTACTTGTGTGTTTCCAGCAATTTGGAGTTCCTGCGATTTTACGGAAGCTAGC	5535	FT	variation	replace(5218,G)	
QY	5383	CGACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTGGCTACAGAGACCTAT	5442	FT	variation	/*tag= b	
DB	5536	CGACAGACGATCTCAAGTGGAGTCTTACCTTTGGCCCTTGGCTACAGAGACCTAT	5595	FT	variation	/standard_name= "Single nucleotide polymorphism"	
QY	5443	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTGAGATCCAGCAGCTCTCAGCA	5502	FT	variation	replace(6065,G)	
DB	5596	CTGTTTGTGACCCACTTCAACTCACTCGAAGTAAATTGAGATCCAGCAGCTCTCAGCA	5655	PN	WO200259325-A2.	/standard_name= "Single nucleotide polymorphism"	
QY	5503	GGGACCCCTGCCGAGGCTACCTGGACATCCGAAACCCGCGCTACCTGGGCCCTGCAAT	5562	XX	01-AUG-2002.		
DB	5656	GGGACCCCTGCCGAGGCTACCTGGACATCCGAAACCCGCGCTACCTGGGCCCTGCAAT	5715	XX	20-DEC-2001; 2001WO-US050497.		
QY	5563	TCCTCAGGAGCGATTACTTTGGGCTCTCATACGAGTAAATTAGGGTTCATTTGCTGC	5622	XX	27-DEC-2000; 2000US-0258335P.		
DB	5716	TCCTCAGGAGCGATTACTTTGGGCTCTCATACGAGTAAATTAGGGTTCATTTGCTGC	5775	XX	(LEXI-) LEXICON GENETICS INC.		
QY	5623	AAGGAAACCTCTGTAAGAGTCCGAGCTGAACACACCGGGGCCCGTCCACCTCCGC	5682	XX	Yu X, Miranda M, Fiddle CJ;		
DB	5776	AAGGAAACCTCTGTAAGAGTCCGAGCTGAACACACCGGGGCCCGTCCACCTCCGC	5835	XX	WPI; 2002-599796/64.		
QY	5683	AGCAGCCCAACAGCGGCGCCACCCAGTCAACAGCAGCATACACGAGCGCTGGCC	5742	XX	P-PSDB; ABB81927.		
DB	5836	AGCAGCCCAACAGCGGCGCCACCCAGTCAACAGCAGCATACACGAGCGCTGGCC	5895	XX	Novel polynucleotide encoding human proteins that are structurally		
QY	5743	TCCAGCCCAACAGCGGCGCCACCCAGTCAACAGCAGCATACACGAGCGCTGGCC	5802	XX	similar to animal kinases, useful for drug screening, diagnosis, in gene		
DB	5896	TCCAGCCCAACAGCGGCGCCACCCAGTCAACAGCAGCATACACGAGCGCTGGCC	5955	XX	therapy of disorders and diseases e.g. cancer and pharmacogenomic		
QY	5803	TACCGGAGGGGGAGCGAGTCCGAGGAGCAAGTCTCTGCGGCCCGCTGGAGGA	5862	XX	Claim 1; Page 37-39; 50pp; English.		
DB	5956	TACCGGAGGGGGAGCGAGTCCGAGGAGCAAGTCTCTGCGGCCCGCTGGAGGA	6015	XX	The invention relates to a novel human protein that shares structural		
QY	5863	GAGAAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGTGTTT	5922	XX	similarity with animal kinases, including serine-threonine kinases,		
DB	6016	GAGAAGTCCCGCGCGGATGCTCAGCAGCGGAGAGCGGTCCCGCGGAGGTGTTT	6075	XX	particularly Citron rho-interacting kinases. The proteins of the		
				CC	invention have neotropic and cytostatic activity. The polynucleotides may		
				CC	have a use in gene therapy. The encoded novel polypeptides are useful for		
				CC	generating antibodies, as reagents in diagnostic assays, for identifying		
				CC	other cellular gene products related to NIP and as reagents in assays for		
				CC	screening for compounds that are useful in the treatment of mental,		
				CC	biological or medical disorders and diseases including cancer. The		
				CC	sequence encodes a novel human kinase of the invention		
				XX	Sequence 6165 BP; 1735 A; 1550 C; 1679 G; 1201 T; 0 U; 0 Other;		
				XX	Query Match 91.9%; Score 5661.4; DB 6; Length 6165;		
				XX	Best Local Similarity 95.8%; Pred. No. 0;		

Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

```
QY 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATTCCTTTTGGATGCTGGTGTCTGCTGAACCCATT 60
Db 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATTCCTTTTGGATGCTGGTGTCTGCTGAACCCATT 60

QY 61 GCCAGCGGGCCTCCAGGCTGAATCTGTTCTCCAGGGNAACCAACCCCTTTATGACTCAA 120
Db 61 GCCAGCGGGCCTCCAGGCTGAATCTGTTCTCCAGGGNAACCAACCCCTTTATGACTCAA 120

QY 121 CAGCAGATGTCCTCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180
Db 121 CAGCAGATGTCCTCTCTTTCCGAGAGGGATATTAGATGCCCTCTTTGTTCTCTTTGAA 180

QY 181 GAATGAGTCAGCTGCTCTGATGAAGATTAAGCAGTGAGCACTTTGTCCGGAAGTAT 240
Db 181 GAATGAGTCAGCTGCTCTGATGAAGATTAAGCAGTGAGCACTTTGTCCGGAAGTAT 240

QY 241 TCCGACACCATAGCTGAGTTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 300
Db 241 TCCGACACCATAGCTGAGTTTACAGGAGCTCCAGCCTTCGGCAAGGACTTCGAAGTCAGA 300

QY 301 AGTCTTTAGTGTGCTGCTCACTTTGCTGAAGTGCAGGTGTTAAGAGAAAACCAACCGGG 360
Db 301 AGTCTTTAGTGTGCTGCTCACTTTGCTGAAGTGCAGGTGTTAAGAGAAAACCAACCGGG 360

QY 361 GACATCTATGCTATGAAGTGTGAAGAGAGAGGCTTTATTTGSCCCAGGAGCAGGTTTCA 420
Db 361 GACATCTATGCTATGAAGTGTGAAGAGAGAGGCTTTATTTGSCCCAGGAGCAGGTTTCA 420

QY 421 TTTTGTGAGAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480
Db 421 TTTTGTGAGAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGGATCCCCCAATTA 480

QY 481 CAGATGCTCTTCAGACAAAATCACTTTATCTGATGAGGAGATATCAGCCTGGAGGG 540
Db 481 CAGATGCTCTTCAGACAAAATCACTTTATCTGATGAGGAGATATCAGCCTGGAGGG 540

QY 541 GACTTCTGCTCACTTTTGAATGATATGAGGACCAAGTTAGATGAAAACCTGATACAGTTT 600
Db 541 GACTTCTGCTCACTTTTGAATGATATGAGGACCAAGTTAGATGAAAACCTGATACAGTTT 600

QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGTTTCATCTGATGGGATACGTGCATCGA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTCAAGGTTTCATCTGATGGGATACGTGCATCGA 660

QY 661 GACATCAAGCCTGAGAACATCTCTGTTGACCGCAGAGACATCAAGCTGGTGGATTTT 720
Db 661 GACATCAAGCCTGAGAACATCTCTGTTGACCGCAGAGACATCAAGCTGGTGGATTTT 720

QY 721 GGATCTGCCGGAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780
Db 721 GGATCTGCCGGAAAATGAATTCAAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780

QY 781 CCAGATTATAGTCTCCTCAAGTGTGACTGTGATGAGGAGGATGGAAGGACCTTAC 840
Db 781 CCAGATTATAGTCTCCTCAAGTGTGACTGTGATGAGGAGGATGGAAGGACCTTAC 840

QY 841 GGCCTGGACTGTGACTGTGGTCASTGGGCGTCAITTCCTATGAGATGATTTATGGGAGA 900
Db 841 GGCCTGGACTGTGACTGTGGTCASTGGGCGTCAITTCCTATGAGATGATTTATGGGAGA 900

QY 901 TCCCCCTTCGAGAGGGAACCTCTGCCAGAACCTTCATTAATGAATTAATTCAGCGG 960
Db 901 TCCCCCTTCGAGAGGGAACCTCTGCCAGAACCTTCATTAATGAATTAATTCAGCGG 960

QY 961 TTTTGTGAAATTTCCAGATCACCCCAAGTGAGCAGTACTTTCTTGATCTGATTCAGAGC 1020
Db 961 TTTTGTGAAATTTCCAGATCACCCCAAGTGAGCAGTACTTTCTTGATCTGATTCAGAGC 1020

QY 1021 TTGTTGTGCGGCAGAAAAGAGAGCTGAAGTTTGAAGTCTTTTGTGCTGCATCTTTCTTC 1080
Db 1021 TTGTTGTGCGGCAGAAAAGAGAGCTGAAGTTTGAAGTCTTTTGTGCTGCATCTTTCTTC 1080
```

```
QY 1081 TCTAAATTTGACTGGAAACAACATTCGTAACTCTCTCTCCCTCTTCTGTTCCACCCCTCAAG 1140
Db 1081 TCTAAATTTGACTGGAAACAACATTCGTAACTCTCTCTCCCTCTTCTGTTCCACCCCTCAAG 1140

QY 1141 TCTGACGATGACACCTCCAAATTTTGTATGAAACAGAGAAGAATTCGTGGGTTTCATCTCT 1200
Db 1141 TCTGACGATGACACCTCCAAATTTTGTATGAAACAGAGAAGAATTCGTGGGTTTCATCTCT 1200

QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTTCG 1260
Db 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGTGGGTTTTTCG 1260

QY 1261 TACAGCAAGGCACTGGGGATTCCTTGGTAGATCTGAGTCTGTTGTCTGGGCTGAGCTCC 1320
Db 1261 TACAGCAAGGCACTGGGGATTCCTTGGTAGATCTGAGTCTGTTGTCTGGGCTGAGCTCC 1320

QY 1321 CCTGCAAGACTAGCTCCATGAAAAGAACTTCTCTCAAAAGCAAGAGCTCAAGAC 1380
Db 1321 CCTGCAAGACTAGCTCCATGAAAAGAACTTCTCTCAAAAGCAAGAGCTCAAGAC 1380

QY 1381 TCTCAGGCAAGTGTCAAGATCGAGCAGAAAATGACCCGGTTTACATCGGAGAGTGTCA 1440
Db 1381 TCTCAGGCAAGTGTCAAGATCGAGCAGAAAATGACCCGGTTTACATCGGAGAGTGTCA 1440

QY 1441 GAGTGGAGGCTGTGCTTAGTCAGAAGGAGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500
Db 1441 GAGTGGAGGCTGTGCTTAGTCAGAAGGAGTGGAGCTGAAGCCTCTGAGACTCAGAGA 1500

QY 1501 TCCCTCTCGAGCAGGACCTTTGCTTACCTACATCAAGAAATGACAGTAGCTTAAAGCGAAGT 1560
Db 1501 TCCCTCTCGAGCAGGACCTTTGCTTACCTACATCAAGAAATGACAGTAGCTTAAAGCGAAGT 1560

QY 1561 TTGAGCAAGCAGGATGGAGGTCTCCAGAGAGATGACAAAGCACTTCGACCTTCTCCAT 1620
Db 1561 TTGAGCAAGCAGGATGGAGGTCTCCAGAGAGATGACAAAGCACTTCGACCTTCTCCAT 1620

QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACACAGGCTCAA 1680
Db 1621 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACACAGGCTCAA 1680

QY 1681 GTGAAAGAAATGAGTTGATGAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740
Db 1681 GTGAAAGAAATGAGTTGATGAATCAAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740

QY 1741 CGGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAGATTTCAAG 1800
Db 1741 CGGAGTGTCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAGATTTCAAG 1800

QY 1801 CGGAAAGCGACAGAATCTCAGCATAAACTGTGAAAGCTTAAGGATCAAGGGAAGCCTGAA 1860
Db 1801 CGGAAAGCGACAGAATCTCAGCATAAACTGTGAAAGCTTAAGGATCAAGGGAAGCCTGAA 1860

QY 1861 GTGGGAAATATGCAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
Db 1861 GTGGGAAATATGCAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920

QY 1921 CTCCAAGAGAAACTGGAGAAAGGCT----- 1944
Db 1921 CTCCAAGAGAAACTGGAGAAAGGCT----- 1944

QY 1945 -----GCAAGGAGCGAGCGAGGGAGCTGGAGAAAGCTTGAGAAAGCTTGAGAAAGCTGCTG 1980
Db 1945 -----GCAAGGAGCGAGCGAGGGAGCTTGAGAAAGCTTGAGAAAGCTTGAGAAAGCTGCTG 1980

QY 1981 AATATCGCGAGGCAAGAGGAGCGAGCGAGGGAGCTGGAGAAAGCTTGAGAAAGCTGAGAG 2040
Db 1981 AATATCGCGAGGCAAGAGGAGCGAGCGAGGGAGCTGGAGAAAGCTTGAGAAAGCTGAGAG 2040

QY 1993 GATTCTTTCTGAAGCATCAGAAAGAGCTGGTGGAAAGCTGAGGAAAGCTGAGAAAGCTTCTCTG 2052
Db 1993 GATTCTTTCTGAAGCATCAGAAAGAGCTGGTGGAAAGCTGAGGAAAGCTGAGAAAGCTTCTCTG 2052

QY 2053 GAGAACAGGTAAAGAGACTTAGAGACCATGAGCGCTAGAGAAAGCAGACTGAAGAGTAC 2112
Db 2053 GAGAACAGGTAAAGAGACTTAGAGACCATGAGCGCTAGAGAAAGCAGACTGAAGAGTAC 2112
```

2113 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2172  
2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAAATTTCTGGAGCTCGAAGAG 2220  
2173 AAAATCTGGAGGCCCAAGTCTCAGCCAGACCTTAGAAGTGCACCTGAAACAGAAAGAG 2232  
2221 AAAATCTGGAGGCCCAAGTCTCAGCCAGACCTTAGAAGTGCACCTGAAACAGAAAGAG 2280  
2233 CAGCACTATGAGAAAGATTAAGTGTGGCAATCAGATTAAGAAAGACCTGCTGAC 2292  
2281 CAGCACTATGAGAAAGATTAAGTGTGGCAATCAGATTAAGAAAGACCTGCTGAC 2340  
2293 AAGGAGACATCTGAGAAACATGATGCAGACACGAGGAGGAGGCCCATGAGAGGGGCAAA 2352  
2341 AAGGAGACATCTGAGAAACATGATGCAGACACGAGGAGGAGGCCCATGAGAGGGGCAAA 2400  
2353 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAA 2412  
2401 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGGATTCCAGATCAGATCCCTGGAA 2460  
2413 CAGAGGATTGTGAACTGTCTGAAGCCAAATAAATTTGCAGCAATAGCAGTCTTTTACC 2472  
2461 CAGAGGATTGTGAACTGTCTGAAGCCAAATAAATTTGCAGCAATAGCAGTCTTTTACC 2520  
2473 CAAAGGAACATGAAGGCCCAAGAGGAGATGATTTCTGAACTCAGGCCAAGAAAATTTTAC 2532  
2521 CAAAGGAACATGAAGGCCCAAGAGGAGATGATTTCTGAACTCAGGCCAAGAAAATTTTAC 2580  
2533 CTGGAGACACAGCTGGAGGTTGGAGGCCCAAGCCGAAATCTGAGGAGCAGCTGGAG 2592  
2581 CTGGAGACACAGCTGGAGGTTGGAGGCCCAAGCCGAAATCTGAGGAGCAGCTGGAG 2640  
2593 AAGATCAGGCCAACAGACACACAGTGAAGAAATCGGCTCTGGAATCGGAACTGGAGCAAGATTG 2652  
2641 AAGATCAGGCCAACAGACACACAGTGAAGAAATCGGCTCTGGAATCGGAACTGGAGCAAGATTG 2700  
2653 CGGAGGTGAGTCTAGAGCAGAGGAGCAGAACTGGAGCTCAGGCCAGCTCAGAG 2712  
2701 CGGAGGTGAGTCTAGAGCAGAGGAGCAGAACTGGAGCTCAGGCCAGCTCAGAG 2760  
2713 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTCAGAGCCCTGAGGCTGCAAGGCG 2772  
2761 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTCAGAGCCCTGAGGCTGCAAGGCG 2820  
2773 GCCTTGAGAGCAGGTTCCGCGAGGAGCAGAGTGAAGAGACACAGAGAGCT 2832  
2821 GCCTTGAGAGCAGGTTCCGCGAGGAGCAGAGTGAAGAGACACAGAGAGCT 2880  
2833 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAATCCAGCGCAATTTGATGCT 2892  
2881 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAATCCAGCGCAATTTGATGCT 2940  
2893 CTTCTGAACAGCTGTACTGTAAATCAGAGCCTCGAGGAGCAGCTAAACAGCTGACCGAG 2952  
2941 CTTCTGAACAGCTGTACTGTAAATCAGAGCCTCGAGGAGCAGCTAAACAGCTGACCGAG 3000  
2953 GACACGCTGAATCAACACCAAACTTCTACTGTCACAACTCAGATGAGGCTTCT 3012  
3001 GACACGCTGAATCAACACCAAACTTCTACTGTCACAACTCAGATGAGGCTTCT 3060  
3013 GGGCCCAACGACAGATTTGTAACCTCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3072  
3061 GGGCCCAACGACAGATTTGTAACCTCGAAGTGAAGTGGACCATCTCCGCCGGGAGATC 3120  
3073 ACGGAACGAGAGTGAAGTTACAGCCAGAACCAACAGTGGAGCTCTGAGACCAAG 3132  
3121 ACGGAACGAGAGTGAAGTTACAGCCAGAACCAACAGTGGAGCTCTGAGACCAAG 3180  
3133 TGCACCATGCTGGAGGAACAGGTCATGGAATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3192  
3181 TGCACCATGCTGGAGGAACAGGTCATGGAATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3240  
3193 AAAGAGCGGAGTGGAGGGCTGGAGAGGCTCTGGGTGATGAGAAATCCCAAGTTTGAG 3252

3241 AAAGAGCGGAGTGGAGGGCTGGAGAGGCTCTGGGTGATGAGAAATCCCAAGTTTGAG 3300  
3253 TGTCCGGTTTCGAGAGCTGCAGAGAAATGCTGACACCGAGAAAACAGAGCAGGGGAGAGCC 3312  
3301 TGTCCGGTTTCGAGAGCTGCAGAGAAATGCTGACACCGAGAAAACAGAGCAGGGGAGAGCC 3360  
3313 GATCAGCGGATCAACCGAGTCTCGCCAGGTGGAGTGGAGTGGCAGTGAAGAGCAAGAGCT 3372  
3361 GATCAGCGGATCAACCGAGTCTCGCCAGGTGGAGTGGAGTGGCAGTGAAGAGCAAGAGCT 3420  
3373 GAGATTCTCGCTCTGACAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGAGCCTCTCT 3432  
3421 GAGATTCTCGCTCTGACAGGCTCTCAAGAGCAGAGCTGAAGCCCGAGAGCCTCTCT 3480  
3433 GACAAGCTCAATGACCTGGAGAGAAAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3492  
3481 GACAAGCTCAATGACCTGGAGAGAAAGCATGCTATGCTTGAATGAATGCCCGAAGCTTA 3540  
3493 CAGCAGAAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3552  
3541 CAGCAGAAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3600  
3553 TTACAGCAGCAGATGGAGCTGCAAGAAAATCAATTTTCGCTCTGACTCAAGGACTGCAA 3612  
3601 TTACAGCAGCAGATGGAGCTGCAAGAAAATCAATTTTCGCTCTGACTCAAGGACTGCAA 3660  
3613 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGTGAATTTGGAGTATCAGCTG 3672  
3661 GAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGTGAATTTGGAGTATCAGCTG 3720  
3673 GAAACCAATTCAGTCTCTATTTCTGAAAGGTGAAATGGAAGCACTATTTCTCAA 3732  
3721 GAAACCAATTCAGTCTCTATTTCTGAAAGGTGAAATGGAAGCACTATTTCTCAA 3780  
3733 CAAACCAATTCATTTGTAAGTGAAGCACTATTTCTCAAAGTGAAGCACTATTTCTCAA 3792  
3781 CAAACCAATTCATTTGTAAGTGAAGCACTATTTCTCAAAGTGAAGCACTATTTCTCAA 3838  
3793 GGTTATTTAGTCACGGAAAGGACCTGCTTACCCACACAGGTTCTCTGCAAGTAC 3852  
3839 -----AGGTTCTCTGCAAGTAC 3855  
3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGTCTGCTGTGCAGAGCTAGAGGAAGCC 3912  
3856 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGTCTGCTGTGCAGAGCTAGAGGAAGCC 3915  
3913 CTTCAGAAAGCCCGATCGAGCTCCGCTCCGCGGGAGGAAGTGCCTCCACCGCAAGCA 3972  
3916 CTTCAGAAAGCCCGATCGAGCTCCGCTCCGCGGGAGGAAGTGCCTCCACCGCAAGCA 3975  
3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGCGAGGAGCAGATCGCCATGTCGCCC 4032  
3976 ACGGACCAACCAACCCATCCAGCCAGCCACCGCGAGGAGCAGATCGCCATGTCGCCC 4035  
4033 ATCGTGGCGTCCGAGAGCAGCCAGCCAGTGCATGAGCTGTGGCCCGCCATCCAGC 4092  
4036 ATCGTGGCGTCCGAGAGCAGCCAGCCAGTGCATGAGCTGTGGCCCGCCATCCAGC 4095  
4093 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCTTAAAGGAACGATGCAC 4152  
4096 CGCAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGGCTCTTAAAGGAACGATGCAC 4155  
4153 CACAATATTCCTACCGATTCACAGTAGGACTGAACATGGAGGCAACAAGTGTGTGTG 4212  
4156 CACAATATTCCTACCGATTCACAGTAGGACTGAACATGGAGGCAACAAGTGTGTGTG 4215  
4213 TGTCTGAGTACCGTGCACCTTTGGAGCCAGCATCCCAATGCTCTCGAATGTTCAGTGTATG 4272  
4216 TGTCTGAGTACCGTGCACCTTTGGAGCCAGGCATCCCAATGCTCTCGAATGTTCAGTGTATG 4275  
4273 TGTCAACCCAAAGTGTCTCCAGCTGCTTGCAGCCACCTGCGGCTTGCCTGTGAATATGCC 4332

Db 4276 TGTCACCCCAAGTGCTCCAGTGCTTTGCCAGCCACCTCGGGCTTGCTGCTGAATATGCC 4335  
QY 4333 ACACACTTCACCGAGGCTTTCTGCCGTGACAAATGAATCCTCCAGGTCTCCAGACCAAG 4392  
Db 4336 ACACACTTCACCGAGGCTTTCTGCCGTGACAAATGAATCCTCCAGGTCTCCAGACCAAG 4395  
QY 4393 GAGCCACAGAGAGCTTGACCTGGAGAGGTGATGAAGTGCACAGGATCAACAAACGA 4452  
Db 4396 GAGCCACAGAGAGCTTGACCTGGAGAGGTGATGAAGTGCACAGGATCAACAAACGA 4455  
QY 4453 GAGACCAAGGCTGGACAGGAGTACATTGCTCGAGGATCAAAAAGTCTCTCATTTAT 4512  
Db 4456 GAGACCAAGGCTGGACAGGAGTACATTGCTCGAGGATCAAAAAGTCTCTCATTTAT 4515  
QY 4513 GACAAATGAAGCAGACAGAGCTGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCTTCC 4572  
Db 4516 GACAAATGAAGCAGACAGAGCTGACAGAGCCGGTGGAGAAATTTGAGCTGTGCCTTCC 4575  
QY 4573 GAGGGGATGTATCTATTCAATGGTGCCTGTTGGTCTTCCGAACTCGCAAAATACAGCCAAA 4632  
Db 4576 GAGGGGATGTATCTATTCAATGGTGCCTGTTGGTCTTCCGAACTCGCAAAATACAGCCAAA 4635  
QY 4633 GCA----- 4635  
Db 4636 GCAGATGTCCCATACATACTGAAGATGGAATCTCACCCGCACACCACCTGCTGGCCCGGG 4695  
QY 4636 ----- 4635  
Db 4696 AGAACCTCTACTTGTAGTCTCCAGCTTCCCTGTGACAAACAGCGGTGGGTACCCGCTTA 4755  
QY 4636 -----GAAAAAGCAGAAAGCTGATCTAAACTG 4662  
Db 4756 GAATCAGTTGTGCGAGGTGGGAGAGTTTCTAGGGAANAAGACAGAGCTGATGCTAAACTG 4815  
QY 4663 CTTGGAACTCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAACTGCAAGCTG 4722  
Db 4816 CTTGGAACTCCCTGCTGAACTGGAAGGTGATGACCGTCTAGACATGAACTGCAAGCTG 4875  
QY 4723 CCCTTCAGTGACAGGTGTTGTTGGGCACCGCAGGAAGGCTCTAGCCCTGAAATGTC 4782  
Db 4876 CCCTTCAGTGACAGGTGTTGTTGGGCACCGCAGGAAGGCTCTAGCCCTGAAATGTC 4935  
QY 4783 TTGAAAACTCCCTAAACCATGTCCTCCAGAAATGGAGCAGTCTCCAAATTTATATTATC 4842  
Db 4936 TTGAAAACTCCCTAAACCATGTCCTCCAGAAATGGAGCAGTCTCCAAATTTATATTATC 4995  
QY 4843 AAGGACCTGGAGAGTACTCATGATAGCAGAGAGAGCGGCACCTGTGCTCTGTGGAC 4902  
Db 4996 AAGGACCTGGAGAGTACTCATGATAGCAGAGAGAGCGGCACCTGTGCTCTGTGGAC 5055  
QY 4903 GTGAAGAAAGTGAACAGTCCCTGCCCCAGTCCCCACCTGCCTGCCCGACCGACATCTCA 4962  
Db 5056 GTGAAGAAAGTGAACAGTCCCTGCCCCAGTCCCCACCTGCCTGCCCGACCGACATCTCA 5115  
QY 4963 CCCAACATTTTGAAGCTGTGAAGGCTGCCACTGTTTGGGGCAGGCAAGATTGAGAAC 5022  
Db 5116 CCCAACATTTTGAAGCTGTGAAGGCTGCCACTGTTTGGGGCAGGCAAGATTGAGAAC 5175  
QY 5023 GGGCTCTGATCTGTGSCACCTATGCCAGCAAGTGGTCAATTCCTCGCTTACAAACGAAAC 5082  
Db 5176 GGGCTCTGATCTGTGSCACCTATGCCAGCAAGTGGTCAATTCCTCGCTTACAAACGAAAC 5235  
QY 5083 CTCAGCAAAATCTGCTCCGAAAGAGATAGAGACTCAGAGCCCTGAGCTGTATCCAC 5142  
Db 5236 CTCAGCAAAATCTGCTCCGAAAGAGATAGAGACTCAGAGCCCTGAGCTGTATCCAC 5295  
QY 5143 TTCACCAATACAGTATCTCTATTGGAACCAATAAATTCAGAAATCGACATGAAGCAG 5202  
Db 5296 TTCACCAATACAGTATCTCTATTGGAACCAATAAATTCAGAAATCGACATGAAGCAG 5355  
QY 5203 TCACGCTCGAGAAATCTCTGATGAAGATGACCAATCTCTGGCACCTGCTGTGTTGCC 5262  
Db 5356 TCACGCTCGAGAAATCTCTGATGAAGATGACCAATCTCTGGCACCTGCTGTGTTGCC 5415

## RESULT 7

AAS06701

ID AAS06701 standard; cDNA; 6159 BP.

XX AAS06701;

XX AAS06701;

DT 12-SEP-2001 (first entry)

XX Polynucleotide sequence encoding human protein kinase #1.

DE Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
KW metabolic disorder; immune related disease; neurological disorder;  
KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
KW reproductive disorder; gene therapy; ss.

OS Homo sapiens.

XX WO200138503-A2.

XX

XX



Db	1501	AGATCCCTCCTGAGCAGGACCTTGCTACCTACATCACAGATGCAGTAGCTTAAAGCGA	1560
Qy	1558	AGTTTGGAGCAAGCA CGGATGAGAGTGTCCAGAGGATGACAAAGCATGCGAGTTTCTC	1617
Db	1561	AGTTTGGAGCAAGCA CGGATGAGAGTGTCCAGAGGATGACAAAGCATGCGAGTTTCTC	1620
Qy	1618	CATGATATCAGAGAGCAGAGCCCGAAGCTCAAAGAAATCAAAGAGCAGGAGTACCAAGGCT	1677
Db	1621	CATGATATCAGAGAGCAGAGCCCGAAGCTCAAAGAAATCAAAGAGCAGGAGTACCAAGGCT	1680
Qy	1678	CAAGTGGAGAAATGAGTTGATGATGAATCAGTTTGGAGAGATCTTTCTCAGCAAGA	1737
Db	1681	CAAGTGGAGAAATGAGTTGATGATGAATCAGTTTGGAGAGATCTTTCTCAGCAAGA	1740
Qy	1738	AGACGAGTGTATCTTACGAATCTGAGCTGAGAGTCTCGGCTTCTCTCTGAGAAATTC	1797
Db	1741	AGACGAGTGTATCTTACGAATCTGAGCTGAGAGTCTCGGCTTCTCTCTGAGAAATTC	1800
Qy	1798	AAGCGGAAAGCGACAGAAATGTACGATAAATCTGTTGAAGGCTAAGGATCAAGGGAGCCT	1857
Db	1801	AAGCGGAAAGCGACAGAAATGTACGATAAATCTGTTGAAGGCTAAGGATCAAGGGAGCCT	1860
Qy	1858	GAGTGGGAGATATCGGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAG	1917
Db	1861	GAGTGGGAGATATCGGAACTGGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAG	1920
Qy	1918	GAGCTCCAAAGAGAAA CTGGAGAAGGCT-----	1944
Db	1921	GAGCTCCAAAGAGAAA CTGGAGAAGGCTGTAAGAGCCAGCAGGAGCCACCGAGCTGCTG	1980
Qy	1945	-----GCAAAGAGCGAGCGCGAGGAGCTGGAGAGCTGGAGAGCTGAGAAACCGA	1989
Db	1981	CAGAAATATCCGCGAGGCAAGAGCGAGCGGAGGAGCTGGAGAGCTGGAGAAACCGA	2040
Qy	1990	GAGGATTTCTTCAAGGCTCAGAAAGAGCTGTGTGAGAGCTGAGGAGCGCGCCATTTCT	2049
Db	2041	GAGGATTTCTTCAAGGCTCAGAAAGAGCTGTGTGAGAGCTGAGGAGCGCGCCATTTCT	2100
Qy	2050	CTGGAGAA CAAGGTAAAGAGATAGAGACCATGGAGCGTAGAGAAAACAGACTGAAGAT	2109
Db	2101	CTGGAGAA CAAGGTAAAGAGATAGAGACCATGGAGCGTAGAGAAAACAGACTGAAGAT	2160
Qy	2110	GACATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGAGCTCGAA	2169
Db	2161	GACATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTCTGAGCTCGAA	2220
Qy	2170	GAGAAACATCGGGAGGCCCAAGTCTCAGCCCGACACCTAGAAAGTGCACCTGAAACAGAAA	2229
Db	2221	GAGAAACATCGGGAGGCCCAAGTCTCAGCCCGACACCTAGAAAGTGCACCTGAAACAGAAA	2280
Qy	2230	GAGCAGCACTATGAGGAAAGATTAAGTTGTTGGACAAATCAGATTAAGAAAGACCTGGCT	2289
Db	2281	GAGCAGCACTATGAGGAAAGATTAAGTTGTTGGACAAATCAGATTAAGAAAGACCTGGCT	2340
Qy	2290	GACAGGAGACACTGGAGAACTATGACAGACACGAGGAGGAGGCCCATGAGAGGGC	2349
Db	2341	GACAGGAGACACTGGAGAACTATGACAGACACGAGGAGGAGGCCCATGAGAGGGC	2400
Qy	2350	AAAAATTTCTCAGCGAACAGAGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTG	2409
Db	2401	AAAAATTTCTCAGCGAACAGAGCGATGATCAATGCTATGGATTCCAAGATCAGATCCCTG	2460
Qy	2410	GACAGAGATTTGGAACTGTCTGAGCCATAAATCTGAGCAAAATAGCAGTCTTTTT	2469
Db	2461	GACAGAGATTTGGAACTGTCTGAGCCATAAATCTGAGCAAAATAGCAGTCTTTTT	2520
Qy	2470	ACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAAATTT	2529
Db	2521	ACCCAAAGGAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAAATTT	2580
Qy	2530	TACCTTGGAGACACAGGCTGGAGTTTGGAGGCCCAAGAACCGAAACTGGAGGAGCAGCTG	2589
Db	2581	TACCTTGGAGACACAGGCTGGAGTTTGGAGGCCCAAGAACCGAAACTGGAGGAGCAGCTG	2640
Qy	2590	GAGAAAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTGCTGGAACTGGAGACAAGA	2649
Db	2641	GAGAAAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTGCTGGAACTGGAGACAAGA	2700
Qy	2650	TTGCGGGAGGTCAGTCTAGAGCAGAGGAGCAGAAACTGGAGCTCAAGGCCAGCTCACA	2709
Db	2701	TTGCGGGAGGTCAGTCTAGAGCAGAGGAGCAGAAACTGGAGCTCAAGGCCAGCTCACA	2760
Qy	2710	GAGCTACAGCTCTCCCTCAGAGGCGGAGTCAACAGTTTGAAGGCCCTCAGGCTGCACGG	2769
Db	2761	GAGCTACAGCTCTCCCTCAGAGGCGGAGTCAACAGTTTGAAGGCCCTCAGGCTGCACGG	2820
Qy	2770	GCGGCCCTGGAGAGCCAGCTTCGCCAGCGGAGCAGAGCTGGAGAGACCAACAGCAGAA	2829
Db	2821	GCGGCCCTGGAGAGCCAGCTTCGCCAGCGGAGCAGAGCTGGAGAGACCAACAGCAGAA	2880
Qy	2830	GCTGAAGAGGAGATCCAGGCCACTCACGGCACATAGAGATGAAATCCAGCGCAAAATTTGAT	2889
Db	2881	GCTGAAGAGGAGATCCAGGCCACTCACGGCACATAGAGATGAAATCCAGCGCAAAATTTGAT	2940
Qy	2890	GCTCTTCTGTAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCTTAAACAGCTGACCT	2949
Db	2941	GCTCTTCTGTAACAGCTGTACTGTAAATCACAGACCTGGAGGAGCTTAAACAGCTGACCT	3000
Qy	2950	GAGGACAAACGCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCT	3009
Db	3001	GAGGACAAACGCTGAACTCAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCT	3060
Qy	3010	TCTGGGCGCAACGACGAGATGTTTACAACTGCGAAAGTGAAGTGGACCTATCTCCGCGGGAG	3069
Db	3061	TCTGGGCGCAACGACGAGATGTTTACAACTGCGAAAGTGAAGTGGACCTATCTCCGCGGGAG	3120
Qy	3070	ATCACGAAACGAGAGATGACGCTTACCAGCCAGAGCAAAAGATGGAGGCTCTGAAGACC	3129
Db	3121	ATCACGAAACGAGAGATGACGCTTACCAGCCAGAGCAAAAGATGGAGGCTCTGAAGACC	3180
Qy	3130	ACGTGCAACCATGCTGGAGGAAACAGGTATGGAATTTGGAGGCCCTTAAACGATGAGCTGCTA	3189
Db	3181	ACGTGCAACCATGCTGGAGGAAACAGGTATGGAATTTGGAGGCCCTTAAACGATGAGCTGCTA	3240
Qy	3190	GAAAAAGAGCGGCTGGAGAGGCTTGAGAGGCTCTCGGTGATGAGAAATCCAGTTT	3249
Db	3241	GAAAAAGAGCGGCTGGAGAGGCTTGAGAGGCTCTCGGTGATGAGAAATCCAGTTT	3300
Qy	3250	GAGTGTCCGGTTTCGAGAGCTGACAGAAATGCTGACACCCAGAGAAAACAGAGCGGCGAGA	3309
Db	3301	GAGTGTCCGGTTTCGAGAGCTGACAGAAATGCTGACACCCAGAGAAAACAGAGCGGCGAGA	3360
Qy	3310	GCGATCAGCGGATCACCGAGTCTCGCCAGTGGAGCTGGAGTGGAGTGAAGGAGCACAAG	3369
Db	3361	GCGATCAGCGGATCACCGAGTCTCGCCAGTGGAGCTGGAGTGGAGTGAAGGAGCACAAG	3420
Qy	3370	GCTGAGATTTCTGCTCTGACGAGGCTCTCAAGAGCAGAGCTGAGGCGGAGGCTC	3429
Db	3421	GCTGAGATTTCTGCTCTGACGAGGCTCTCAAGAGCAGAGCTGAGGCGGAGGCTC	3480
Qy	3430	TCTGCAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAATGAAATGAAATGAAATG	3489
Db	3481	TCTGCAAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAAATGAAATGAAATGAAATG	3540
Qy	3490	TTACAGCAGAGCTGGAGACTGAAACAGAGCTCAAAACAGAGGCTTCTGAAAGAGCAGGCC	3549
Db	3541	TTACAGCAGAGCTGGAGACTGAAACAGAGCTCAAAACAGAGGCTTCTGAAAGAGCAGGCC	3600
Qy	3550	AAATTTACAGCAGAGATGACCTGCAAGAAAATTCACATTTTCCGCTCTGACTCAAGGACTG	3609
Db	3601	AAATTTACAGCAGAGATGACCTGCAAGAAAATTCACATTTTCCGCTCTGACTCAAGGACTG	3660
Qy	3610	CAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATCTGGAGTATCAG	3669
Db	3661	CAAGAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGAGTGAATCTGGAGTATCAG	3720





Dd 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCTTCGGCAAAAGGACTTCGAAGTCAGA 300  
Qy 301 AGCTCTGTAGTTGTGGTCACTTTGCTGAAGTGCAGTGTGAAGAGAGAAAGCAACCGGG 360  
Dd 301 AGCTCTGTAGTTGTGGTCACTTTGCTGAAGTGCAGTGTGAAGAGAGAAAGCAACCGGG 360  
Qy 361 GACATCTATGCTATGAAAGTGATGAAGAGAGAGGCTTTATTTGGCCGAGGAGAGTTTCA 420  
Dd 361 GACATCTATGCTATGAAAGTGATGAAGAGAGAGGCTTTATTTGGCCGAGGAGAGTTTCA 420  
Qy 421 TTTTGTGAGAGAGCGGACATATTTATCTGAGAGACAGAGCCCGTGGATCCGCCAATTA 480  
Dd 421 TTTTGTGAGAGAGCGGACATATTTATCTGAGAGACAGAGCCCGTGGATCCGCCAATTA 480  
Qy 481 CAGTATGCTCTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGGCTCGAGGG 540  
Dd 481 CAGTATGCTCTTCAGGACAAAATACCTTTATCTGATGGAGGAATATCAGGCTCGAGGG 540  
Qy 541 GACTTGTCTGCTTTGATATGATGAGGACAGCTTAGATGAAGAACCTGATACAGTTT 600  
Dd 541 GACTTGTCTGCTTTGATATGATGAGGACAGCTTAGATGAAGAACCTGATACAGTTT 600  
Qy 601 TACTAGCTGAGCTGATTTTGGCTGTTTACAGGTTTCACTGATGGGATACGTCATCGA 660  
Dd 601 TACTAGCTGAGCTGATTTTGGCTGTTTACAGGTTTCACTGATGGGATACGTCATCGA 660  
Qy 661 GACATCAAGCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGTGGATTTT 720  
Dd 661 GACATCAAGCTGAGAACATTTCTGTTGACCGACAGGACACATCAAGCTGTGGATTTT 720  
Qy 721 GGATCTGCCGCGAAATGAATTCACACAGATGGTGAATGCCAAATCCCGATTTGGACC 780  
Dd 721 GGATCTGCCGCGAAATGAATTCACACAA --GGTGAATGCCAAATCCCGATTTGGACC 777  
Qy 781 CCAGATTACATGCTCTGAAAGTCTGACTGTGATGAACGGGATGGAAGGACCTTAC 840  
Dd 778 CCAGATTACATGCTCTGAAAGTCTGACTGTGATGAACGGGATGGAAGGACCTTAC 837  
Qy 841 GGCTGGACGTGACTGTGGTCACTGGGGGTGATTCCTATGAGATGATTTATGGGAGA 900  
Dd 838 GGCTGGACGTGACTGTGGTCACTGGGGGTGATTCCTATGAGATGATTTATGGGAGA 897  
Qy 901 TCCCTCTTCGACAGGGAACCTCTGCAGAACCTTCAATAAATATGAATTTCCAGCGG 960  
Dd 898 TCCCTCTTCGACAGGGAACCTCTGCAGAACCTTCAATAAATATGAATTTCCAGCGG 957  
Qy 961 TTTTGAATTTCCAGTGAACCCAAAGTGAGCAGTGACTTTCTTGATCTGATTCGAAGC 1020  
Dd 958 TTTTGAATTTCCAGTGAACCCAAAGTGAGCAGTGACTTTCTTGATCTGATTCGAAGC 1017  
Qy 1021 TTCTGTGGGCGGACAGAGACTGAAGTTTGAAGGTCCTTGTGCGCATCCTTTCTTC 1080  
Dd 1018 TTGTGTGGGCGGACAGAGACTGAAGTTTGAAGGTCCTTGTGCGCATCCTTTCTTC 1077  
Qy 1081 TCTAAATTTGACTGGAACCAATTCGTAACTCTCTCCCTCTCTCCCTCTCCCTCTCAAG 1140  
Dd 1078 TCTAAATTTGACTGGAACCAATTCGTAACTCTCTCCCTCTCTCCCTCTCCCTCTCAAG 1137  
Qy 1141 TCTGAGATGACACTTCCAAATTTTGTATGACACAGAGAAATTCGTGGGTTTCATCCTCT 1200  
Dd 1138 TCTGAGATGACACTTCCAAATTTTGTATGACACAGAGAAATTCGTGGGTTTCATCCTCT 1197  
Qy 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260  
Dd 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG 1257  
Qy 1261 TACAGAGGCACTGGGATTTCTTGTAGATCTGAGTCTTGTGTGCGGGTCTGAGCTCC 1320  
Dd 1258 TACAGAGGCACTGGGATTTCTTGTAGATCTGAGTCTTGTGTGCGGGTCTGAGCTCC 1317  
Qy 1321 CCGTGCAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1380  
Dd 1318 CCGTGCAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAAGAC 1377

Qy 1381 TCTCAGGACAAGTGTCAAGATGAGCAGAGAAATGACCCGGTTTACATCGAGAGTGTCA 1440  
Dd 1378 TCTCAGGACAAGTGTCAAGATGAGCAGAGAAATGACCCGGTTTACATCGAGAGTGTCA 1437  
Qy 1441 GAGTGGAGGCTGTGCTTAGTCAGAGAGGTTGAGCTGAAGCCCTCTGAGACTCAGAGA 1500  
Dd 1438 GAGTGGAGGCTGTGCTTAGTCAGAGAGGTTGAGCTGAAGCCCTCTGAGACTCAGAGA 1497  
Qy 1501 TCCCTCTCTGAGCAGGACCTTGTACCTACATCACAGAATGCAGTAGCTTTAAAGCCGAAGT 1560  
Dd 1498 TCCCTCTCTGAGCAGGACCTTGTACCTACATCACAGAATGCAGTAGCTTTAAAGCCGAAGT 1557  
Qy 1561 TTGAGCAAGCACGAGTGGAGGTTCTCCAGAGGATGACAAAAGCACTGCGACTTCTCAT 1620  
Dd 1558 TTGAGCAAGCACGAGTGGAGGTTCTCCAGAGGATGACAAAAGCACTGCGACTTCTCAT 1617  
Qy 1621 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1680  
Dd 1618 GATATCAGAGCAGAGCCGGAAGCTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1677  
Qy 1681 GTGGAAGAAATGAGGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1740  
Dd 1678 GTGGAAGAAATGAGGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAAGA 1737  
Qy 1741 CGGAGTGTCTCTACGAATCTGAGCTGAGAGGTTCTCGGCTTGTCTGAGAAATTCAG 1800  
Dd 1738 CGGAGTGTCTCTACGAATCTGAGCTGAGAGGTTCTCGGCTTGTCTGAGAAATTCAG 1797  
Qy 1801 CGGAAAGCGCAGAAATGTCAGCATAAACTGTTGAAGGCTAAGGATCAAGGAAAGCCTGAA 1860  
Dd 1798 CGGAAAGCGCAGAAATGTCAGCATAAACTGTTGAAGGCTAAGGATCAAGGAAAGCCTGAA 1857  
Qy 1861 GTGGGAGAAATGCGAAAATCTGGAGAAATCAATGCTCAGCAGCAGCTCAAAATTCAGGAG 1920  
Dd 1858 GTGGGAGAAATGCGAAAATCTGGAGAAATCAATGCTCAGCAGCAGCTCAAAATTCAGGAG 1917  
Qy 1921 CTCCAAGAGAAACTGGAGAAAGGCTG----- 1945  
Dd 1918 CTCCAAGAGAAACTGGAGAAAGGCTGTTAAAGCCAGCAGCGGAGCCACCGAGTCTGCAG 1977  
Qy 1946 -----CAAGGAGCGAGCGCGAGGAGCTGGAGAGCTGGAGAACCGAGAG 1992  
Dd 1978 AATATCCCGCAGGCAAGGAGCGAGCGCGAGGAGCTGGAGAAAGCTGCAGAACCGGAG 2037  
Qy 1993 GATTCCTTCTGAAGCATCAGAAAGAGCTGTGGAAGCTCAGGAAAGCCGCCATTTCTG 2052  
Dd 2038 GATTCCTTCTGAAGCATCAGAAAGAGCTGTGGAAGCTGAGGAAAGCCGCCATTTCTG 2097  
Qy 2053 GAGAAACAAGGTAAAGAGACTAGAGACCATGGAGCCATGAGAGAAACAGACTGAAGGATGAC 2112  
Dd 2098 GAGAAACAAGGTAAAGAGACTAGAGACCATGGAGCCATGAGAGAAACAGACTGAAGGATGAC 2157  
Qy 2113 ATCCAGACAAATCCCAACAGATCCAGCATGCTGATATAAATTTCTGAGCTCGAAGAG 2172  
Dd 2158 ATCCAGACAAATCCCAACAGATCCAGCATGCTGATATAAATTTCTGAGCTCGAAGAG 2217  
Qy 2173 AAACATCGGAGGCGCCCAAGTCTCAGCCCGACACCTAGAAGTGCACCTGAAACAGAAAGAG 2232  
Dd 2218 AAACATCGGAGGCGCCCAAGTCTCAGCCCGACACCTAGAAGTGCACCTGAAACAGAAAGAG 2277  
Qy 2233 CAGCACTATGAGGAAAGATTAAGTGTGGACATCAGATAAAGAAAGACCTGGCTGAC 2292  
Dd 2278 CAGCACTATGAGGAAAGATTAAGTGTGGACATCAGATAAAGAAAGACCTGGCTGAC 2337  
Qy 2293 AAGGAGACACTGGAGAAACATGATGCAGAGACACAGAGGAGGCGCCATGAGAAAGGCGCAA 2352  
Dd 2338 AAGGAGACACTGGAGAAACATGATGCAGAGACACAGAGGAGGCGCCATGAGAAAGGCGCAA 2397  
Qy 2353 ATTCTCAGGACACAGAGGCGGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2412  
Dd 2398 ATTCTCAGGACACAGAGGCGGATGATCAATGCTATGATTTCCAGATCAGATCCCTGGAA 2457

QY 2413 CAGAGATTGTGGAACTGTCTGAAGCCAAATAAACTTGCAGCAAAATAGCAGTCTTTTACC 2472  
DB 2458 CAGAGATTGTGGAACTGTCTGAAGCCAAATAAACTTGCAGCAAAATAGCAGTCTTTTACC 2517  
QY 2473 CAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAAATTTTAC 2532  
DB 2518 CAAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAAATTTTAC 2577  
QY 2533 CTGGAGACACAGGCTGGGAGTTGGAGGCCCAAGACCGAAACTGGAGGAGCAGCTGGAG 2592  
DB 2578 CTGGAGACACAGGCTGGGAGTTGGAGGCCCAAGACCGAAACTGGAGGAGCAGCTGGAG 2637  
QY 2593 AAGATCAGCCACCAAGACCAACACAGTGAAGAATTCGCTCTCGTGAACCTGAGACCAAGATTG 2652  
DB 2638 AAGATCAGCCACCAAGACCAACACAGTGAAGAATTCGCTCTCGTGAACCTGAGACCAAGATTG 2697  
QY 2653 CGGGAGTCACTAGACACAGGAGGACAGAACTGGAGCTCAAGCCCGAGCTCAGAG 2712  
DB 2698 CGGGAGTCACTAGACACAGGAGGACAGAACTGGAGCTCAAGCCCGAGCTCAGAG 2757  
QY 2713 CTACAGCTCTCCTCTGAGGAGCGGAGTCAACAGTTGACAGCCCTGACAGGCTGACGGGCG 2772  
DB 2758 CTACAGCTCTCCTCTGAGGAGCGGAGTCAACAGTTGACAGCCCTGACGCTGACGGGCG 2817  
QY 2773 GGCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGAGACT 2832  
DB 2818 GGCCTGGAGAGCCAGCTTCGCCAGGCGAAGACAGAGCTGGAAGAGACACACAGAGACT 2877  
QY 2833 GAAGAGGAGATCCAGGCACTCAGGCACATAGAGATGAATCAGCGCAAAATTTGATGCT 2892  
DB 2878 GAAGAGGAGATCCAGGCACTCAGGCACATAGAGATGAATCAGCGCAAAATTTGATGCT 2937  
QY 2893 CTTCTGTAACAGTGTACTATATCAAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2952  
DB 2938 CTTCTGTAACAGTGTACTATATCAAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2997  
QY 2953 GACAGCTGAACTCAACACCAAACTTCTACTTTGTCGAAAGTGAAGTGAACCTCCGCGGAGATC 3012  
DB 2998 GACAGCTGAACTCAACACCAAACTTCTACTTTGTCGAAACAACTCGATGAGGCTTCT 3057  
QY 3013 GCGGCCAACGAGAGATTGACAACTGCGAAGTGAAGTGAACCTCCGCGGAGATC 3072  
DB 3058 GCGGCCAACGAGAGATTGACAACTGCGAAGTGAAGTGAACCTCCGCGGAGATC 3117  
QY 3073 ACGGAACGAGATGAGCTTACGAGCAGAGAGCAAAAGATGAGGCTCTGAAGACCAAG 3132  
DB 3118 ACGGAACGAGATGAGCTTACGAGCAGAGAGCAAAAGATGAGGCTCTGAAGACCAAG 3177  
QY 3133 TGCACCATCTCGAGGAAACAGTCTATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3192  
DB 3178 TGCACCATCTCGAGGAAACAGTCTATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3237  
QY 3193 AAAGAGCGGCAATGGAGGCTGGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3252  
DB 3238 AAAGAGCGGCAATGGAGGCTGGAGGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3297  
QY 3253 TGTCCGGTTCGAGAGCTGAGAGATGCTGACACCGAGAAACAGAGCGGCGAGAGCC 3312  
DB 3298 TGTCCGGTTCGAGAGCTGAGAGATGCTGACACCGAGAAACAGAGCGGCGAGAGCC 3357  
QY 3313 GATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAGGCT 3372  
DB 3358 GATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAGGAGCACAGGCT 3417  
QY 3373 GAGATTCTCGTCTGAGAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGGCTCTCT 3432  
DB 3418 GAGATTCTCGTCTGAGAGGCTCTCAAGAGCAGAGCTGAAGCCGAGAGGCTCTCT 3477  
QY 3433 GACAGCTCAATGACCTGAGAGAGAGCATGCTATGCTTTGAAATGAATGCCCGAGGCTTA 3492  
DB 3478 GACAGCTCAATGACCTGAGAGAGAGCATGCTATGCTTTGAAATGAATGCCCGAGGCTTA 3537  
QY 3493 CAGCAGAGCTGGAGACTGAACAGAGCTCAAAACAGAGGCTTCTGGAAAGACAGCCAAA 3552

DB 3558 CAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAGAGCAGCAACAA 3597  
QY 3553 TTACAGAGCAGATGGAGCTTCAGAGAAAATCACAATTTTCGTCTGACTCAAGGACTGCAA 3612  
DB 3598 TTACAGAGCAGATGGAGCTTCAGAGAAAATCACAATTTTCGTCTGACTCAAGGACTGCAA 3657  
QY 3613 GAAGCTCTAGATCGGCTGATCTACTCAAGACAGAAAAGTGAAGTCTTGGAGTATCAGCTG 3672  
DB 3658 GAAGCTCTAGATCGGCTGATCTACTCAAGACAGAAAAGTGAAGTCTTGGAGTATCAGCTG 3717  
QY 3673 GAAAACATTCAGGTTCTTATTTCTCATGAAAAGTGAAGTGAAGCACTATTTCTCAA 3732  
DB 3718 GAAAACATTCAGGTTCTTATTTCTCATGAAAAGTGAAGTGAAGCACTATTTCTCAA 3777  
QY 3733 CAACCCAACTCATTTGATTTCTGCAAGCCAAATGCAACCTGCTTAAAGAAAAG 3792  
DB 3778 CAACCCAACTCATTTGATTTCTGCAAGCCAAATGCAACCTGCTTAAAGAAAAG - 3835  
QY 3793 GGTTTATTTAGTCACGAGAAAGAGGACCTGCTTTTACCCACACAGGTTCTCTCAGTAC 3852  
DB 3836 -----AGGTGCTCTCAGTAC 3852  
QY 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTTAGAGAAAGCC 3912  
DB 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTTAGAGAAAGCC 3912  
QY 3913 CTTTCAAGAGACCCGATCGAGCTCCGCTCCGCGGAGGAGGCTGCCACCGAAAGCA 3972  
DB 3913 CTTTCAAGAGACCCGATCGAGCTCCGCTCCGCGGAGGAGGCTGCCACCGAAAGCA 3972  
QY 3973 ACGAACCAACCAACCAATCCAGCCAGCAACCGAGGAGGAGAGTGGCATATTCGCGC 4032  
DB 3973 ACGAACCAACCAACCAATCCAGCCAGCAACCGAGGAGGAGAGTGGCATATTCGCGC 4032  
QY 4033 ATCGTCGGTCCGAGAGCACCCAGCTGCGCATGAGCTGCTGGCCCGCCCATCCAGC 4092  
DB 4033 ATCGTCGGTCCGAGAGCACCCAGCTGCGCATGAGCTGCTGGCCCGCCCATCCAGC 4092  
QY 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTCGGGCTCTTAAAGAACCGATCAC 4152  
DB 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAAATTTAGTCGGGCTCTTAAAGAACCGATCAC 4152  
QY 4153 CACAATATTCCTACCCGATTCAGCTAGGACTGNAACATGCGAGGCCACAAAGTGTGCTGTG 4212  
DB 4153 CACAATATTCCTACCCGATTCAGCTAGGACTGNAACATGCGAGGCCACAAAGTGTGCTGTG 4212  
QY 4213 TGTCTGATACCGTGCACTTTGGACGCCAGCATCCAAATGTCTCGAATGTCTCAGGTGATG 4272  
DB 4213 TGTCTGATACCGTGCACTTTGGACGCCAGCATCCAAATGTCTCGAATGTCTCAGGTGATG 4272  
QY 4273 TGTCAACCCCAAGTGTCTCCAGTGTCCAGGCCACCTGCGGCTTGCCTGTGAATATGCC 4332  
DB 4273 TGTCAACCCCAAGTGTCTCCAGTGTCCAGGCCACCTGCGGCTTGCCTGTGAATATGCC 4332  
QY 4333 ACACATTCACCGAGGCTTCTCGCGTGACAAATGAACCTCCCGAGGTCTCCAGACCAAG 4392  
DB 4333 ACACATTCACCGAGGCTTCTCGCGTGACAAATGAACCTCCCGAGGTCTCCAGACCAAG 4392  
QY 4393 GAGCCACGAGCAGCTTGCACCTTGAAGGCTGATGAAGTGGCCAGGAATAACAAACGA 4452  
DB 4393 GAGCCACGAGCAGCTTGCACCTTGAAGGCTGATGAAGTGGCCAGGAATAACAAACGA 4452  
QY 4453 GGACAGCAAGGCTGGGACAGAGATCATTTGCTTGGAGGATCAAAAGTCTCATTTAT 4512  
DB 4453 GGACAGCAAGGCTGGGACAGAGATCATTTGCTTGGAGGATCAAAAGTCTCATTTAT 4512  
QY 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCCGTGGAGAAATTTGAGCTGTGCTTCCC 4572  
DB 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCCGTGGAGAAATTTGAGCTGTGCTTCCC 4572  
QY 4573 GACGGGATGTATCTATTCTATGCTGCTTGGTTCGAACTCGCAATAACAGCCAAA 4632

Db 4573 GACGGGATGATATCTATTCATGTGCGTGGTGGTCTTCGAACTCCGAAATACAGCCAAA 4632  
 Qy 4633 GCA----- 4635  
 Db 4633 GCAGATGCCATACATACATAGAGATGAATCTCACCCGACACACCTGCTGGCCGGG 4692  
 Qy 4636 ----- 4635  
 Db 4693 AGAACCTCTACTTGTAGTCCAGCTTCCCTGACAAAACAGCGTGGGTCAACCGCTTA 4752  
 Qy 4636 -----GAAAAAGCAGAAGCTGATGCTAACTG 4662  
 Db 4753 GAATCAGTTGTCGACGAGTGGAGAGTTTCTAGGAAAAGCAGAAGCTGATGCTAACTG 4812  
 Qy 4663 CTTGGAAATCCCTGCTGTAACCTGGAAGTGTATGACCGTCTAGACATGAATCGACGCTG 4722  
 Db 4813 CTTGGAAATCCCTGCTGTAACCTGGAAGTGTATGACCGTCTAGACATGAATCGACGCTG 4872  
 Qy 4723 CCCTTCAGTGACACAGTGTGTTGGTGGGACCGAGGAAGGGCTCTACGCCCTGAATGTC 4782  
 Db 4873 CCCTTCAGTGACACAGTGTGTTGGTGGGACCGAGGAAGGGCTCTACGCCCTGAATGTC 4932  
 Qy 4783 TTGAAAACTCCCTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4842  
 Db 4933 TTGAAAACTCCCTAACCCATGTCACAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4992  
 Qy 4843 AAGGACCTGGAGAGTACTACTCATGATAGCAGAGAGAGGGGCACTGTCTGTGGAC 4902  
 Db 4993 AAGGACCTGGAGAGTACTACTCATGATAGCAGGTAAGAGGGGCACTGTCTGTGGAC 5052  
 Qy 4903 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGCCCGACATCTCA 4962  
 Db 5053 GTGAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCCAGCCCGACATCTCA 5112  
 Qy 4963 CCACAACTTTTGAAGTGTCAAGGCTGCACCTTTGTTGGGCGAGGAAAGATTGAGAAC 5022  
 Db 5113 CCACAACTTTTGAAGTGTCAAGGCTGCACCTTTGTTGGGCGAGGAAAGATTGAGAAC 5172  
 Qy 5023 GGGCTCTGCATCTGTGAGGCGATGCCAGCAAGTGTCTATCTCCGCTACACAGAAAC 5082  
 Db 5173 GGGCTCTGCATCTGTGAGGCGATGCCAGCAAGTGTCTATCTCCGCTACACAGAAAC 5232  
 Qy 5083 CTGACGAAATATGTCATTCGGAAGAGATAGAGACCTCAGAGCCCTGACGCTGTATCCAC 5142  
 Db 5233 CTGACGAAATATGTCATTCGGAAGAGATAGAGACCTCAGAGCCCTGACGCTGTATCCAC 5292  
 Qy 5143 TTCAACAAATACAGTATCCCTCATTTGGAACCAATATATTTCTAGAAATCGACATGAGCAG 5202  
 Db 5293 TTCAACAAATACAGTATCCCTCATTTGGAACCAATATATTTCTAGAAATCGACATGAGCAG 5352  
 Qy 5203 TACACGCTCGAGGAATTCCTGGATAAGATGACCATTCCTTGCCACCTGCTGTGTTGCC 5262  
 Db 5353 TACACGCTCGAGGAATTCCTGGATAAGATGACCATTCCTTGCCACCTGCTGTGTTGCC 5412  
 Qy 5263 GCCTCTTCAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGCAGGCGAGGAG 5322  
 Db 5413 GCCTCTTCAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGCAGGCGAGGAG 5472  
 Qy 5323 GAGTACTTGTGTTTCCACGAATTTGGAGTGTCTGCTGGATTTTACGGAAGAGCGTAGC 5382  
 Db 5473 GAGTACTTGTGTTTCCACGAATTTGGAGTGTCTGCTGGATTTTACGGAAGAGCGTAGC 5532  
 Qy 5383 CGACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGGCTTACAGAGAACCTCTAT 5442  
 Db 5533 CGACAGACGATCTCAAGTGGAGTGCCTTACCTTTGGCCCTTTGGCCCTTACAGAGAACCTCTAT 5592  
 Qy 5443 CTGTTTGTGACCCACATTCACCTCACTCGAAGTAATTTGAGATCCAGGACGCTCTCAGCA 5502  
 Db 5593 CTGTTTGTGACCCACATTCACCTCACTCGAAGTAATTTGAGATCCAGGACGCTCTCAGCA 5652  
 Qy 5503 GGGACCCCTGCCGAGGCTACCTTGGACATCCCGAACCCCGGCTTACTGGGCGCTGCCATT 5562  
 Db 5653 GGGACCCCTGCCGAGGCTACCTTGGACATCCCGAACCCCGGCTTACTGGGCGCTGCCATT 5712

Qy 5563 TCCTCAGAGAGCATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGCTGC 5622  
 Db 5713 TCCTCAGAGAGCATTTACTTTGGCGTCTCTATACAGGATAAATTAAGGTCATTTGCTGC 5772  
 Qy 5623 AAGGAAACCTTCGTGAAGGAGTCCGGCACTGAACACACCGGGCCGTCACCTCCCGC 5682  
 Db 5773 AAGGAAACCTTCGTGAAGGAGTCCGGCACTGAACACACCGGGCCGTCACCTCCCGC 5832  
 Qy 5683 AGCAGCCCCAACAAAGCAGAGCCCAACCCACGTAACAGAGCACATCAACAGCGCTGGCC 5742  
 Db 5833 AGCAGCCCCAACAAAGCAGAGCCCAACCCACGTAACAGAGCACATCAACAGCGCTGGCC 5892  
 Qy 5743 TCCAGCCAGCGCCCGGAGGCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 5802  
 Db 5893 TCCAGCCAGCGCCCGGAGGCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 5952  
 Qy 5803 TACCCGAGGCGCGGAGCCAGCTGCGCAGGACAAAGTCTCTGCGCCGCCCTTGAGCGA 5862  
 Db 5953 TACCCGAGGCGCGGAGCCAGCTGCGCAGGACAAAGTCTCTGCGCCGCCCTTGAGCGA 6012  
 Qy 5863 GAGAAGTCTCCCGCCGCGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGCGTGTGTT 5922  
 Db 6013 GAGAAGTCTCCCGCCGCGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGCGTGTGTT 6072  
 Qy 5923 GAACACAGCAGCAGCGGCGCGCTGCTGCGGAGCCGCTGAGGACCCCGCTGTCCAGGTG 5982  
 Db 6073 GAACACAGCAGCAGCGGCGCGCTGCTGCGGAGCCGCTGAGGACCCCGCTGTCCAGGTG 6132  
 Qy 5983 AACAGGAGAGGCGCAGAGTGCCTCTCAAGTTTTCAGGTTTAACTGTCACT 6037  
 Db 6133 AACAGGAGAGGCGCAGAGTGCCTCTCAAGTTTTCAGTATAAATCTCAGCCAGAAAACCAACTCT 6187

RESULT 9  
 ADA05641  
 ID ADA05641 standard; cDNA; 6189 BP.  
 AC ADA05641;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human NOV1a encoding cDNA SEQ ID NO:1.  
 XX  
 KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective;  
 KW antiparkinsonian; antilipidemic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..6162  
 FT /\*tag= a  
 FT /product= "NOV1a"  
 XX  
 PN WO2003029424-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 PP 02-OCT-2002; 2002WO-US031373.  
 XX  
 PR 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.





Db 3118 ACGGAACGAGATGCGAGCTTACCGCCAGAACGATGAGGCTCTGAGACCCAG 3177  
QY 3133 TGCACCATCTGGAGAAACAGGCTCATGATTTGGAGGCCCTAAACGATGAGCTGTAGAA 3192  
Db 3178 TGCACCATGCTGGAGAACAGCTCATGATTTGGAGGCCCTPAAACGATGAGCTGTAGAA 3237  
QY 3193 AAAGAGCGCATGGGAGCGCTGGAGGAGCGTCTCTGGGTGATGAGAAATCCGAGTTTGAG 3252  
Db 3238 AAAGAGCGCATGGGAGCGCTGGAGGAGCGTCTCTGGGTGATGAGAAATCCGAGTTTGAG 3297  
QY 3253 TGTCTGGTTCGAGAGCTGACAGAGATGCTGGACACCCGAGAAACAGAGCGGCGAGAGCC 3312  
Db 3298 TGTCTGGTTCGAGAGCTGACAGAGATGCTGGACACCCGAGAAACAGAGCGGCGAGAGCC 3357  
QY 3313 GATCAGCGGATCACCGAGTCTCCGAGGTGGAGCTGGAGTGGAGTGAAGAGCACAGGCT 3372  
Db 3358 GATCAGCGGATCACCGAGTCTCCGAGGTGGAGCTGGAGTGGAGTGAAGAGCACAGGCT 3417  
QY 3373 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGCTGAAGCGCGAGAGCTCTCT 3432  
Db 3418 GAGATTCTCGCTCTGAGAGGCTCTCAAGAGCAGAGCTGAAGCGCGAGAGCTCTCT 3477  
QY 3433 GACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCGAAAGCTTA 3492  
Db 3478 GACAGCTCAATGACCTGGAGAGAGCATGCTATGCTTGAATGAATGCCGAAAGCTTA 3537  
QY 3493 CAGCAGAGCTGGAGACTGAAACGAGACTCAACAGAGGCTTCTGGAGAGCAAGCCAAA 3552  
Db 3538 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAA 3597  
QY 3553 TTACAGCAGAGTGGAGCTGAGAGAAATCAATTTCCGCTGACTCAAGGACTGCAA 3612  
Db 3598 TTACAGCAGAGTGGAGCTGAGAGAAATCAATTTCCGCTGACTCAAGGACTGCAA 3657  
QY 3613 GAAGCTCTAGATCGGCTGATCTATGAGAGCAGAGAGAGTGAATGGAGTATCAGCTG 3672  
Db 3658 GAAGCTCTAGATCGGCTGATCTATGAGAGCAGAGAGAGTGAATGGAGTATCAGCTG 3717  
QY 3673 GAAACATTTGAGTCTCTATCTCATGAAAGGCTGAAATGAAAGGCACTATTTCTCAA 3732  
Db 3718 GAAACATTTGAGTCTCTATCTCATGAAAGGCTGAAATGAAAGGCACTATTTCTCAA 3777  
QY 3733 CAAACAAACTCATTTGCTGCAAGCCAAATGAGCAACCTGCTAAAAGAAAAAG 3792  
Db 3778 CAAACAAACTCATTTGCTGCAAGCCAAATGAGCAACCTGCTAAAAGAAAA - 3835  
QY 3793 GGTATTATTGTCGACGGAAGAGGACCTGCTTTTACCACACAGGTTCTCTGAGTAC 3852  
Db 3836 -----AGGTGCTCTGAGTAC 3852  
QY 3853 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3912  
Db 3853 AATGAGCTGAAGTGGCCCTGGAGAGGAGAAAGCTCGCTGTGCAGAGCTAGAGGAAGCC 3912  
QY 3913 CTTTCAAGAACCCGCATCGAGTCCCGTCCGCGGGAGGAAAGCTGCCACCGCAAGCA 3972  
Db 3913 CTTTCAAGAACCCGCATCGAGTCCCGTCCGCGGGAGGAAAGCTGCCACCGCAAGCA 3972  
QY 3973 ACGGACCCACCAACCCATCCACGACAGCCACGCGAGGAGCAGATGCCATGTCGCGC 4032  
Db 3973 ACGGACCCACCAACCCATCCACGACAGCCACGCGAGGAGCAGATGCCATGTCGCGC 4032  
QY 4033 ATCGTGGGTGCGCAGAGCAGCCAGTGGCCATGAGCTGCTGSCCCCGCCATCCAGC 4092  
Db 4033 ATCGTGGGTGCGCAGAGCAGCCAGTGGCCATGAGCTGCTGSCCCCGCCATCCAGC 4092  
QY 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTTCGCGGCTTTAAGGAAACGATGCAC 4152  
Db 4093 CGCAGAAAGGAGTCTTCAACTCCAGAGGAATTTAGTTCGCGGCTTTAAGGAAACGATGCAC 4152  
QY 4153 CACAATATTTCTCACCAGTTTCAACGTAGGACTGAAATGCGAGGCCAAGTGTGCTGTG 4212  
Db 4153 CACAATATTTCTCACCAGTTTCAACGTAGGACTGAAATGCGAGGCCAAGTGTGCTGTG 4212

QY 4213 TGTCTGATACCGTGCATTTTGAAGCCAGGCATCCAAATGTCTCGAATGTCTAGGTGATG 4272  
Db 4213 TGTCTGATACCGTGCATTTTGAAGCCAGGCATCCAAATGTCTAGAATGTCTAGGTGATG 4272  
QY 4273 TGTCAACCCCAAGTGTCTCCAGTCTTGCAGCCAGCAGCTGCGGCTTGCCTGCTGAATATGCC 4332  
Db 4273 TGTCAACCCCAAGTGTCTCCAGTCTTGCAGCCAGCAGCTGCGGCTTGCCTGCTGAATATGCC 4332  
QY 4333 ACACATTTACCCAGGCGCTTCTGCGGTGACAAATGAATCTCCAGAGTCTCCAGACCAAG 4392  
Db 4333 ACACATTTACCCAGGCGCTTCTGCGGTGACAAATGAATCTCCAGAGTCTCCAGACCAAG 4392  
QY 4393 GAGCCAGCAGCAGCTTGCACCTTGAAGGCTGATGAGGTGCCAGGATACAAACGA 4452  
Db 4393 GAGCCAGCAGCAGCTTGCACCTTGAAGGCTGATGAGGTGCCAGGATACAAACGA 4452  
QY 4453 GGAAGCAAGGCTGGGACAGGAAATCATTTGCTCTGAGGGATCAAAAGTCTCATTTAT 4512  
Db 4453 GGAAGCAAGGCTGGGACAGGAAATCATTTGCTCTGAGGGATCAAAAGTCTCATTTAT 4512  
QY 4513 GACATGAAGCAGAGAGCTGACAGAGCCGGTGGAAATTTGAGCTGTGCTTCCC 4572  
Db 4513 GACATGAAGCAGAGAGCTGACAGAGCCGGTGGAAATTTGAGCTGTGCTTCCC 4572  
QY 4573 GACGGGATGTATCTATTCTATGCTGCGTGTGCTTCCGAACTCGCAATACAGCCAAA 4632  
Db 4573 GACGGGATGTATCTATTCTATGCTGCGTGTGCTTCCGAACTCGCAATACAGCCAAA 4632  
QY 4633 GCA ----- 4635  
Db 4633 GCAGATGCCCATACATACTGAGATGGAATCTCACCCGCACACCACTGCTGSCCGGG 4692  
QY 4636 ----- 4635  
Db 4693 AGAACCTCTACTTGTAGCTCCAGCTTCCGTGACAAACAGCGCTGGGTCAACGCTTA 4752  
QY 4636 -----GAAAAAGCAGAAAGCTGATGCTAAAATG 4662  
Db 4753 GAATCAGTTGTCGAGGTGGGAGAGTTTCTAGGAAAAAGCAGAGCTGATGCTAAAATG 4812  
QY 4663 CTTGGAATCTCCTGTGTGAAACTGGAAGGTGATGACCGTCTAGACATGAATGCAAGCTG 4722  
Db 4813 CTTGGAATCTCCTGTGTGAAACTGGAAGGTGATGACCGTCTAGACATGAATGCAAGCTG 4872  
QY 4723 CCTTTCAGTGACACAGGTGTGTTGTTGGGCACCGAGGAAGGCTCTACGCCCTGAAATGTC 4782  
Db 4873 CCTTTCAGTGACACAGGTGTGTTGTTGGGCACCGAGGAAGGCTCTACGCCCTGAAATGTC 4932  
QY 4783 TTGAAAAACTCCCTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4842  
Db 4933 TTGAAAAACTCCCTAACCCATGTCCCAGGAATTTGGAGCAGTCTTCCAAATTTATATTATC 4992  
QY 4843 AAGGACCTGAGAGCTACTCATGATAGCAGGAGAGAGCGGCACCTGTCTTGTGCGAC 4902  
Db 4993 AAGGACCTGAGAGAGCTACTCATGATAGCAGGTGAGAGCGGCACCTGTCTTGTGCGAC 5052  
QY 4903 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCCGACATCTCA 4962  
Db 5053 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCAGCCCGACATCTCA 5112  
QY 4963 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGAGCAGAGATTGAGAAC 5022  
Db 5113 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTTGGGAGCAGAGATTGAGAAC 5172  
QY 5023 GGGCTCTGCTATCTGTGAGCCATGCCAGCAAAAGTCTGCTTCTCCCTACACAGAAAAAC 5082  
Db 5173 GGGCTCTGCTATCTGTGAGCCATGCCAGCAAAAGTCTGCTTCTCCCTACACAGAAAAAC 5232  
QY 5083 CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACTCTCAGAGCCCTGCAGCTGTATCCAC 5142  
Db 5233 CTCAGCAAAATCTGCATCCGGAAGAGATAGAGACTCTCAGAGCCCTGCAGCTGTATCCAC 5292

QY 5143 TTACCAATACAGTATCTCTCATTTGGAACCAATAAATTTCTACGAAATCCGACATGAAGCAG 5202  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5293 TTACCAATACAGTATCTCTCATTTGGAACCAATAAATTTCTACGAAATCCGACATGAAGCAG 5352  
 QY 5203 TACAGGCTCAGGAAATTCCTGGATAAGAAATGACCATTCCTTGGACATCTGCTGTTGTTGCC 5262  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5353 TACAGGCTCAGGAAATTCCTGGATAAGAAATGACCATTCCTTGGACATCTGCTGTTGTTGCC 5412  
 QY 5263 GCCTCTTCCAAACAGCTTCCTCTGCTCAATCGTGCAGGTGAACAGCGAGGCGAGCAGAG 5322  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5413 GCCTCTTCCAAACAGCTTCCTCTGCTCAATCGTGCAGGTGAACAGCGAGGCGAGCAGAG 5472  
 QY 5323 GAGTACTTGCTGTGTTTCCAGCAATTTGGAGTGTTCGTGGAATTCCTTACGAGAACCTAGC 5382  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5473 GAGTACTTGCTGTGTTTCCAGCAATTTGGAGTGTTCGTGGAATTCCTTACGAGAACCTAGC 5532  
 QY 5383 CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGCGCTTTGCTTACAGAGAACCTAT 5442  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5533 CGCACAGACGATCTCAAGTGGAGTCGCTTACCTTTGCGCTTTGCTTACAGAGAACCTAT 5592  
 QY 5443 CTGTTGTGACCACTTCAACTCACTCGAAGTAATTTAGATCCAGCACGCTCTCTCAGCA 5502  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5593 CTGTTGTGACCACTTCAACTCACTCGAAGTAATTTAGATCCAGCACGCTCTCTCAGCA 5652  
 QY 5503 GGGACCCCTGCCCGAGGTACCTGGACATCCGAAACCGCGCTTACCTGGGCGCTGCCATT 5562  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5653 GGGACCCCTGCCCGAGGTACCTGGACATCCGAAACCGCGCTTACCTGGGCGCTGCCATT 5712  
 QY 5563 TCCTCAGGAGCGATTACCTTTGGCTCTCTATACAGGATAAATTAAGGTGCTATTTGCTGC 5622  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5713 TCCTCAGGAGCGATTACCTTTGGCTCTCTATACAGGATAAATTAAGGTGCTATTTGCTGC 5772  
 QY 5623 AAGGAAACCTCTGTAAGAGTCCGCACTGAACACACCGGGGCGCTTCACTCCCGC 5682  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5773 AAGGAAACCTCTGTAAGAGTCCGCACTGAACACACCGGGGCGCTTCACTCCCGC 5832  
 QY 5683 AGCAGCCCCAACAGGAGGAGCCACCCACGTAACAGGACATACCAAGCGGTGGCC 5742  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5833 AGCAGCCCCAACAGGAGGAGCCACCCACGTAACAGGACATACCAAGCGGTGGCC 5892  
 QY 5743 TCCAGCCAGCGCGCCCGAGGAGCCAGCCACCCCGAGAGAGCAAGACACCCACCGC 5802  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5893 TCCAGCCAGCGCGCCCGAGGAGCCAGCCACCCCGAGAGAGCAAGACACCCACCGC 5952  
 QY 5803 TACCGGAGGCGGAGCGAGCTCCAGGACAGTCTCTGGCGCCGCTTGGAGCGA 5862  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 5953 TACCGGAGGCGGAGCGAGCTCCAGGACAGTCTCTGGCGCCGCTTGGAGCGA 6012  
 QY 5863 GAGAAGTCCCGCGCGGATGCTCAGCACGCGAGAGAGCGGTCCCGCGGAGGCTGTTT 5922  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 6013 GAGAAGTCCCGCGCGGATGCTCAGCACGCGAGAGAGCGGTCCCGCGGAGGCTGTTT 6072  
 QY 5923 GAGACAGCAGGAGGCGGCTCTGCGGAGCGGTGAGGACCCCGCTGTCAGGTG 5982  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 6073 GAGACAGCAGGAGGCGGCTCTGCGGAGCGGTGAGGACCCCGCTGTCAGGTG 6132  
 QY 5983 AACAGGGAGAGGCGAGGTGCTCTCAAGTTTTCACCGTTTAAACACTGTCACT 6037  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 6133 AACAGGTGAGGAGGAGGCTCTCAAGTTTTCACCGTTTAAACACTGTCACT 6187

RESULT 10  
 ABS63435  
 ID ABS63435 standard; cDNA; 6201 BP.  
 AC ABS63435;  
 XX  
 XX  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX Human cDNA, homologous to kinases, designated NOV3a.  
 DE  
 DE Human; gene; ss; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;  
 KW Human; gene; ss; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;  
 KW atherosclerosis; diabetes; cell signalling; metabolic pathway;

KW cellular receptor; downstream effector; cancer; gene therapy;  
 KW hypertension; congenital heart defect; aortic stenosis; obesity;  
 KW infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;  
 KW neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;  
 KW haematopoietic disease; scleroderma; fertility; immunogen;  
 KW idiopathic thrombocytopenic purpura; graft versus host disease;  
 KW Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;  
 KW systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;  
 KW stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;  
 KW pain; alcoholism; transgenic.  
 XX Homo sapiens.  
 OS WO200226826-A2.  
 PN 04-APR-2002.  
 PD 27-SEP-2001; 2001WO-US042336.  
 PF 27-SEP-2000; 2000US-0235631P.  
 PR 27-SEP-2000; 2000US-0235633P.  
 PR 27-SEP-2000; 2000US-0235808P.  
 PR 27-SEP-2000; 2000US-0236064P.  
 PR 27-SEP-2000; 2000US-0236065P.  
 PR 27-SEP-2000; 2000US-0236066P.  
 PR 28-SEP-2000; 2000US-0236135P.  
 PR 03-OCT-2000; 2000US-0237434P.  
 PR 05-OCT-2000; 2000US-0238321P.  
 PR 06-OCT-2000; 2000US-0238396P.  
 PR 16-MAR-2001; 2001US-0276657P.  
 PR 31-MAY-2001; 2001US-0294823P.  
 PR 12-JUL-2001; 2001US-0304868P.  
 PR 26-SEP-2001; 2001US-00235631.  
 XX (CURA-) CURAGEN CORP.  
 PA Gerlach VL, Macdougall JR, Smithson G, Millet I, Stone D;  
 PI Gunther E, Ellerman K, Grosse WM, Alsobrook JP, Lepley DM;  
 PI Burgess CB, Padigaru M, Kekuda R, Spytek KA, Leach MD, Shinkets RA;  
 XX WPI; 2002-499860/53.  
 DR P-PSDB; ABG78362.  
 XX Novel isolated NOVX polypeptides and polynucleotides homologous to  
 PT attractin, plexin, papin-like family of proteins, useful for treating  
 PT atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and  
 PT stroke.  
 XX Claim 8; Page 40-42; 308pp; English.  
 CC The invention discloses the isolated human polypeptides and  
 CC polynucleotides encoding them, that have been designated NOVX. The  
 CC polypeptides, polynucleotides and antibodies are useful in treating or  
 CC preventing a NOVX-associated disorder which is cardiomyopathy,  
 CC atherosclerosis and diabetes in a human, where the disorder is related to  
 CC cell signal processing and metabolic pathway modulation. They can also be  
 CC used in determining the presence of, or predisposition to, a disease  
 CC associated with altered levels of the polypeptides and polynucleotides of  
 CC any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for  
 CC identifying an agent that binds to, or that modulates the expression or  
 CC activity of the polypeptide, for identifying an agent which is cellular  
 CC receptor or downstream effector, for treating or preventing a NOVX-  
 CC associated disorder and as a pharmaceutical composition comprising the  
 CC polypeptide, polynucleotide or the antibody. The polypeptides and  
 CC polynucleotides are useful in diagnostic applications (e.g. as a marker  
 CC for cancerous cells or tissue types) where their amounts are assessed, or  
 CC for the manufacture of a medicament (e.g. gene therapy) for treating or  
 CC preventing disorders or syndromes such as hypertension, congenital heart  
 CC defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,  
 CC Alzheimer's disease, Parkinson's disorder, neurodegenerative disorders,  
 CC haemophilia, dyslipidemias, haematopoietic diseases, scleroderma,  
 CC fertility, idiopathic thrombocytopenic purpura, graft versus host

CC	diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune
CC	disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,
CC	allergy, stroke, anxiety, Lesh-Nyhan syndrome, schizophrenia, cerebellar
CC	ataxia, pain and alcoholism. They may also be used as immunogens to
CC	produce antibodies specific for the invention, and as vaccines.
CC	Transgenic cells containing a NOVX expressing construct are useful to
CC	transgenic animals containing a NOVX expressing construct for studying the function and/or
CC	activity of the NOVX proteins and for identifying and/or evaluating a NOVX
CC	modulators of NOVX protein activity. Transgenic cells containing a NOVX
CC	expressing construct are useful to produce non-human transgenic animals
CC	for studying the function and/or activity of the NOVX proteins and for
CC	identifying and/or evaluating modulators of NOVX protein activity. The
CC	sequences presented in AB963431-ABS63444 are the human NOV1-NOV8 cDNAs
XX	
SQ	Sequence 6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;
	Query Match            91.4%; Score 5629.8; DB 6; Length 6201;
	Best Local Similarity 95.7%; Pred. No. 0;
	Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 300
QY	1 ATGTTGAAGTTCAAATATGGACGCCGAATCCTTTGGATCGTGCTGTCTGAACCCATT 60
DB	1 ATGTTGAAGTTCAAATATGGACGCCGAATCCTTTGGATCGTGCTGTCTGAACCCATT 60
QY	61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCCTTTATGACTCAA 120
DB	61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCCTTTATGACTCAA 120
QY	121 CAGCAGATGTCCTCTTTCCCGAAGAAGGATATTAGATCCCTCTTTGTTCTCTTTGAA 180
DB	121 CAGCAGATGTCCTCTTTCCCGAAGAAGGNATTAGATGCCCTCTTTGTTCTCTTTGAA 180
QY	181 GAATGCAGTCAGCCTGCTCTCATGAAGATTAAACAACCTGAGCAACTTTGTCGGAGTAGT 240
DB	181 GAATGCAGTCAGCCTGCTCTCATGAAGATTAAACAACCTGAGCAACTTTGTCGGAGTAGT 240
QY	241 TCGCACACCATAGCTGAGTTACAGAGCTCCAGGCTTCGGCAAGAGACTTCGAAGTCAGA 300
DB	241 TCGCACACCATAGCTGAGTTACAGAGCTCCAGGCTTCGGCAAGAGACTTCGAAGTCAGA 300
QY	301 AGTCTGTAGGTTGTGTCCTCTTTCGTAAGTGTCAGGTGTTAAAGAGAAAAGCAACCGGG 360
DB	301 AGTCTGTAGGTTGTGTCCTCTTTCGTAAGTGTCAGGTGTTAAAGAGAAAAGCAACCGGG 360
QY	361 GACATCTATGCTATGAAGTGAATGAAGAAGGCTTTATGGCCACGAGACAGGTTTCA 420
DB	361 GACATCTATGCTATGAAGTGAATGAAGAAGGCTTTATTTGGCCACGAGACAGGTTTCA 420
QY	421 TTTTTTGAGGAAGCGGAAACATATTCTCGAAGCAACAAGCCGTGATCCCCCAATTA 480
DB	421 TTTTTTGAGGAAGCGGAAACATATTCTCGAAGCACAGCCGTGATCCCCCAATTA 480
QY	481 CAGTATGCGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCTGGAGGG 540
DB	481 CAGTATGCGCTTTTCAGGACAAAAATCACCTTTATCTGATGGAGGAATATCAGCTGGAGGG 540
QY	541 GACTTCTGCTCACTTTTGAATAGATATCAGGACACAGTTAGATGAACACCTGTATACAGTTT 600
DB	541 GACTTCTGCTCACTTTTGAATAGATATCAGGACACAGTTAGATGAACACCTGTATACAGTTT 600
QY	601 TACCTAGCTGAGCTGATTTTGCGCTGTTCCACAGCGTTTCATCTGTATGGGATACGTGCATCGA 660
DB	601 TACCTAGCTGAGCTGATTTTGCGCTGTTCCACAGCGTTTCATCTGTATGGGATACGTGCATCGG 660
QY	661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGATTTT 720
DB	661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGATTTT 720
QY	721 GGATCTGCCGCGAAATGAATTCAAACAGATGGTGAATGCGCAAACTCCCGATTTGGGACC 780
DB	721 GGATCTGCCGCGAAATGAATTCAAACAA ---GGTGAATGCGCAAACTCCCGATTTGGGACC 777
QY	781 CCAGATTATATGGCTCCTCGAAGTGCTGATGTGATGAACGGGGATGGAAGAGGCACCTAC 840



4033 ATCGTGGGTCGCCAGAGCACCAGCCAGTGCCTATGAGCCCTGCTGGCCCGCCCATCCAGC 4092  
Db |||||  
4033 ATCGTGGGTCGCCAGAGCACCAGCCAGTGCCTATGAGCCCTGCTGGCCCGCCCATCCAGC 4092  
Qy |||||  
4093 CCAGAAAGGAGTCTTCACTCCAGAGAAATTAGTCGGCGCTCTTAAGAAACGATCCAC 4152  
Db |||||  
4093 CCAGAAAGGAGTCTTCACTCCAGAGAAATTAGTCGGCGCTCTTAAGAAACGATCCAC 4152  
Qy |||||  
4153 CACAATATTCCTCACCGGATCAACAGTAGAGTGAATGAGCAGCAGCACAAGTGTGCTGTG 4212  
Db |||||  
4153 CACAATATTCCTCACCGGATCAACAGTAGAGTGAATGAGCAGCAGCACAAGTGTGCTGTG 4212  
Qy |||||  
4213 TGTCTGGATACCGTGCATCTTTGGAGCGCCAGGATCCAAATGTCTCGAATGTCAAGTGATG 4272  
Db |||||  
4213 TGTCTGGATACCGTGCATCTTTGGAGCGCCAGGATCCAAATGTCTAGAAATGTCAAGTGATG 4272  
Qy |||||  
4273 TGTCAACCCAGTGTCTCAAGTGTCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4332  
Db |||||  
4273 TGTCAACCCAGTGTCTCAAGTGTCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4332  
Qy |||||  
4333 ACACATCTTACCGAGGCTCTTCCGCTGACAAATGAATGAACTCCCAAGTCTCCAGACCAAG 4392  
Db |||||  
4333 ACACATCTTACCGAGGCTCTTCCGCTGACAAATGAATGAACTCCCAAGTCTCCAGACCAAG 4392  
Qy |||||  
4393 GAGCCAGCAGCAGTCTGACCTGGAAGGCTGATGAAGTGCCTGAGGATCAAAAGTCTCAATAT 4452  
Db |||||  
4393 GAGCCAGCAGCAGTCTGACCTGGAAGGCTGATGAAGTGCCTGAGGATCAAAAGTCTCAATAT 4452  
Qy |||||  
4453 GACAGCAAGGCTGGGACAGGAGTACATGTCCTGGAGGATCAAAAGTCTCAATAT 4512  
Db |||||  
4453 GACAGCAAGGCTGGGACAGGAGTACATGTCCTGGAGGATCAAAAGTCTCAATAT 4512  
Qy |||||  
4513 GACATGAAGCAGAGAGTGGAGCAGAGGCGGTGGAGAAATTTGAGCTGTGCTTCC 4572  
Db |||||  
4513 GACATGAAGCAGAGAGTGGAGCAGAGGCGGTGGAGAAATTTGAGCTGTGCTTCC 4572  
Qy |||||  
4573 GAGGGATGTATCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4632  
Db |||||  
4573 GAGGGATGTATCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4632  
Qy |||||  
4633 GCA----- 4635  
Db |||||  
4633 GCAGATGTCCTCATACATAGTGAATCTCACCGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4692  
Qy |||||  
4636 ----- 4635  
Db |||||  
4693 AGAACCTCTACTTGCTAGCTCCAGCTTCCCTGACAAACAGCGGTGGGTACCCGCTTA 4752  
Qy |||||  
4636 ----- 4662  
Db |||||  
4753 GAATCAGTTGTCGAGGTGGAGAGTTTCTAGGGAAGAGCAGAGCTGATGCTAACTG 4812  
Qy |||||  
4663 CTTGGAAACTCCCTGCTGAACTGGAGGTGATGACCTGATAGATGAACTGACGCTG 4722  
Db |||||  
4813 CTTGGAAACTCCCTGCTGAACTGGAGGTGATGACCTGATAGATGAACTGACGCTG 4872  
Qy |||||  
4723 CTTTCAGTGACAGGTGTTGTTGGGACCGGAGAGGCTCTACGCCCTGATGTC 4782  
Db |||||  
4873 CTTTCAGTGACAGGTGTTGTTGGGACCGGAGAGGCTCTACGCCCTGATGTC 4932  
Qy |||||  
4783 TTGAAAACTCCCTAACCCATGTCAGGAAATGGAGAGTCTTCCAAATTTATATTATC 4842  
Db |||||  
4933 TTGAAAACTCCCTAACCCATGTCAGGAAATGGAGAGTCTTCCAAATTTATATTATC 4992  
Qy |||||  
4843 AAGGACTGGAGAGCTACTCATGATAGCAGGAGAGGCGGCTGCTGTTGGAC 4902  
Db |||||  
4993 AAGGACTGGAGAGCTACTCATGATAGCAGGAGAGGCGGCTGCTGTTGGAC 5052  
Qy |||||  
4903 GTGAAGAAAGTGAAGTCCCTGCGCCAGTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4962  
Db |||||  
5053 GTGAAGAAAGTGAAGTCCCTGCGCCAGTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5112  
Qy |||||  
4963 CCACAATTTTGAAGTGTCAAGGGCTGCCACTTGTGTTGGGCGAGGAGATGGAAC 5022

RESULT 11

5113 CCACAATTTTGAAGTGTCAAGGGCTGCCCTGTTTGGGCGAGCAAGATTGGAAC 5172  
Qy |||||  
5023 GGGCTCTGCTATCTGTCAGCAGTCCAGCAAGTGTCTATCTCCGCTACAAAGAAAAC 5082  
Db |||||  
5173 GGGCTCTGCTATCTGTCAGCAGTCCAGCAAGTGTCTATCTCCGCTACAAAGAAAAC 5232  
Qy |||||  
5083 CTGAGCAATATCTGCTATCGGAAAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5142  
Db |||||  
5233 CTGAGCAATATCTGCTATCGGAAAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5292  
Qy |||||  
5143 TTCAACAATATACAGTATCTCTATTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5202  
Db |||||  
5293 TTCAACAATATACAGTATCTCTATTGGAACCAATAAATTTCTACGAAATCGACATGAAGCAG 5352  
Qy |||||  
5203 TACAAGCTCGAGAAATCTCTGGAATAAGTACATCTCTTGGGCACTGCTGTGTTGCC 5262  
Db |||||  
5353 TACAAGCTCGAGAAATCTCTGGAATAAGTACATCTCTTGGGCACTGCTGTGTTGCC 5412  
Qy |||||  
5263 GGGCTCTGCAACAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGAGGCGAGGAGAG 5322  
Db |||||  
5413 GGGCTCTTCCAAAGCTTCCCTGCTCAATCGTCAGGTGAACAGCGAGGCGAGGAGAG 5472  
Qy |||||  
5323 GAGTACTTGTGCTGTTTCCAGAAATTTGGAGTGTCTCTGGAATTTTACGGAAGAGCTGAGC 5382  
Db |||||  
5473 GAGTACTTGTGCTGTTTCCAGAAATTTGGAGTGTCTCTGGAATTTTACGGAAGAGCTGAGC 5532  
Qy |||||  
5383 CGACAGAGATCTCAAGTGGAGTGTCTTACCTTTGGCCCTTGGCTTACAGAGAACCTAT 5442  
Db |||||  
5533 CGACAGAGATCTCAAGTGGAGTGTCTTACCTTTGGCCCTTGGCTTACAGAGAACCTAT 5592  
Qy |||||  
5443 CTGTTTCTGACCACTTCAACTCACTCGAAGTAAATTTGAGATCCAGGACAGCTCTCTACAGCA 5502  
Db |||||  
5593 CTGTTTCTGACCACTTCAACTCACTCGAAGTAAATTTGAGATCCAGGACAGCTCTCTACAGCA 5652  
Qy |||||  
5503 GGGACCCCTGCGGAGCGTACTCTGGAATCCCGAAACCCGCGCTACTCTGGGCGCTGCCATT 5562  
Db |||||  
5653 GGGACCCCTGCGGAGCGTACTCTGGAATCCCGAAACCCGCGCTACTCTGGGCGCTGCCATT 5712  
Qy |||||  
5563 TCCTCAGAGCGATTTACTTTGGGCTCTCATACAGGATAAATTTAAGGGTCAATTTGCTGC 5622  
Db |||||  
5713 TCCTCAGAGCGATTTACTTTGGGCTCTCATACAGGATAAATTTAAGGGTCAATTTGCTGC 5772  
Qy |||||  
5623 AAGGAAAACCTCTGTAAGGAGTCCGGCACTGAACACCAACCGGGGCGCTGCCACTCCCGC 5682  
Db |||||  
5773 AAGGAAAACCTCTGTAAGGAGTCCGGCACTGAACACCAACCGGGGCGCTGCCACTCCCGC 5832  
Qy |||||  
5683 AGCAGCCCAACAGCGAGGCGCCACCCAGTACAAAGCAGACATCAACAGCGCGTGGCC 5742  
Db |||||  
5833 AGCAGCCCAACAGCGAGGCGCCACCCAGTACAAAGCAGACATCAACAGCGCGTGGCC 5892  
Qy |||||  
5743 TCCAGCCAGCGCGCGCCGAGGCGCCAGCCACCCGCGAGAGCCAAAGCACACCCACCGC 5802  
Db |||||  
5893 TCCAGCCAGCGCGCGCCGAGGCGCCAGCCACCCGCGAGAGCCAAAGCACACCCACCGC 5952  
Qy |||||  
5803 TACCGGAGGGGCGGACCGAGCTCGGAGGGAAGAATCTCTGCGCCGCGCTGGAGGGA 5862  
Db |||||  
5953 TACCGGAGGGGCGGACCGAGCTCGGAGGGAAGAATCTCTGCGCCGCGCTGGAGGGA 6012  
Qy |||||  
5863 GAGAAGTCCCGCGCGGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGGAGCTGTT 5922  
Db |||||  
6013 GAGAAGTCCCGCGCGGATGCTCAGCACCGGAGAGAGCGGTCCCGCGGAGGAGCTGTT 6072  
Qy |||||  
5923 GAAGACAGCAGCAGGGGCGGCTGCTTGGGAGCGCTGAGGAGCCCGCTGTGCCAGGTG 5982  
Db |||||  
6073 GAAGACAGCAGCAGGGGCGGCTGCTTGGGAGCGCTGAGGAGCCCGCTGTGCCAGGTG 6132  
Qy |||||  
5983 AACAGGGAAGAGGGA 5999  
Db |||||  
6133 AACAGGGAAGAGGGA 6149

ADA05653

ID ADA05653 standard; cDNA; 6201 BP.

XX AC

ADA05653;

XX DT

06-NOV-2003 (first entry)

XX XX

Human NOV1g encoding cDNA SEQ ID NO:13.

XX KW

human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

XX KW immunomodulator; cytosolic; nootropic; neuroprotective;

XX KW antiparkinsonian; antilipemic; gene therapy; human disease;

XX KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

XX KW immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.

XX OS

Homo sapiens.

XX XX

Key Location/Qualifiers

FT CDS 1..6201

FT /\*tag= a

FT /product= "NOV1g"

XX XX

W02003029424-A2.

XX XX

10-APR-2003.

XX XX

02-OCT-2002; 2002WO-US031373.

XX PR

02-OCT-2001; 2001US-0326483P.

XX PR

05-OCT-2001; 2001US-0327435P.

XX PR

05-OCT-2001; 2001US-0327449P.

XX PR

08-OCT-2001; 2001US-0327517P.

XX PR

09-OCT-2001; 2001US-0328029P.

XX PR

09-OCT-2001; 2001US-0328044P.

XX PR

09-OCT-2001; 2001US-0328056P.

XX PR

12-OCT-2001; 2001US-0328849P.

XX PR

15-OCT-2001; 2001US-0329414P.

XX PR

17-OCT-2001; 2001US-0330142P.

XX PR

18-OCT-2001; 2001US-0330309P.

XX PR

22-OCT-2001; 2001US-0341058P.

XX PR

24-OCT-2001; 2001US-0339266P.

XX PR

24-OCT-2001; 2001US-0343629P.

XX PR

29-OCT-2001; 2001US-0349575P.

XX PR

01-NOV-2001; 2001US-0346357P.

XX PR

17-APR-2002; 2002US-0373260P.

XX PR

19-APR-2002; 2002US-0373815P.

XX PR

19-APR-2002; 2002US-0373817P.

XX PR

19-APR-2002; 2002US-0373826P.

XX PR

22-APR-2002; 2002US-0373884P.

XX PR

16-MAY-2002; 2002US-0381037P.

XX PR

16-MAY-2002; 2002US-0381038P.

XX PR

16-MAY-2002; 2002US-0381042P.

XX PR

28-MAY-2002; 2002US-0381642P.

XX PR

29-MAY-2002; 2002US-0383656P.

XX PR

29-MAY-2002; 2002US-0383831P.

XX PR

25-JUN-2002; 2002US-0391335P.

XX PR

01-OCT-2002; 2002US-00262511.

XX PR

(CURA-) CURAGEN CORP.

XX PA

Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

XX PI

Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

XX PI

Ort T, Gorwan L, Zehusen BD, Anderson DW, Zhong M, Catterton E;

XX PI

Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;

XX PI

Shinkets RA, Rothenberg ME, Leach WD, Agee ML, Berghs C, Dipippo VA;

XX PI

Eisen AD, Gangolli EA, Rieger DK, Spaderna SK;

XX PI

WPI; 2003-381626/36.

XX DR

P-PSDB; ADA05654.

XX DR

New NOVX polypeptides and nucleic acids, useful for diagnosing,

XX PT

PT

PT

PT

XX

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;

Query Match 91.4%; Score 5629.8; DB 7; Length 6201;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 5929; Conservative 0; Mismatches 22; Indels 246; Gaps 4;

QY 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATCTGCTGCTGTAACCAATT 60

DB 1 ATGTTGAAGTTCAAATATGAGCGCGGAATCCTTTGGATCTGCTGCTGTAACCAATT 60

QY 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCCAGGGGAAACCACTTTATGACTCAA 120

DB 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCCAGGGGAAACCACTTTATGACTCAA 120

QY 121 CAGCAGATGCTCTCTCTTTCCAGAGAGGATATTAGATGCCCTTTGTTCTTTTGA 180

DB 121 CAGCAGATGCTCTCTCTTTCCAGAGAGGATATTAGATGCCCTTTGTTCTTTTGA 180

QY 181 GAATGAGTCAGCTCTCTCTGATGAAGATTAGCAGCTGAGCAACTTTTCGGAGAGTAT 240

DB 181 GAATGAGTCAGCTCTCTCTGATGAAGATTAGCAGCTGAGCAACTTTTCGGAGAGTAT 240

QY 241 TCCGACACCATAGCTGAGTACAGGAGCTCCAGCCTTCGCGAAGACTTCGAGTCACA 300

DB 241 TCCGACACCATAGCTGAGTACAGGAGCTCCAGCCTTCGCGAAGACTTCGAGTCACA 300

QY 301 AGTCTTTAGGTTGTGCTACCTTTGCTGAAGTCAGGTGTTAAGAGAGAAACCAACCGG 360

DB 301 AGTCTTTAGGTTGTGCTACCTTTGCTGAAGTCAGGTGTTAAGAGAGAAACCAACCGG 360

QY 361 GACATCTATGCTATGAAGTCAGTGAAGAGAGGCTTTTATTCGCCAGGAGGTTTCA 420

DB 361 GACATCTATGCTATGAAGTCAGTGAAGAGAGGCTTTTATTCGCCAGGAGGTTTCA 420

preventing or treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or dyslipidemia, and in chromosome mapping, tissue typing or pharmacogenomics.

Claim 20; Page 104-105; 586pp; English.

The present invention describes NOVX proteins, where X can be 1 to 55 (e.g. NOV1). Also described: (1) a composition comprising a polypeptide described above and a carrier; (2) a kit comprising, in one or more containers, the composition described above; (3) an isolated nucleic acid molecule which encodes a NOVX protein of the invention; (4) a vector comprising the nucleic acid molecule described above; (5) a cell comprising the above vector; (6) an antibody that immunospecifically binds to the polypeptide described above; (7) methods for determining the presence or amount of the above polypeptide or nucleic acid molecule in a sample; (8) methods for determining the presence of or predisposition to a disease associated with altered levels of expression of the above polypeptide or nucleic acid molecule in a first mammalian subject; (9) a method of identifying an agent that binds to the polypeptide described above; (10) a method for identifying a potential therapeutic agent for use in treating a pathology that is related to an aberrant expression or aberrant physiological interactions of the polypeptide; (11) a method of screening for a modulator of activity of or of latency or predisposition to a pathology associated with the polypeptide; (12) a method for modulating the activity of the polypeptide described above; (13) methods of treating or preventing a pathology associated with the above polypeptide in a mammal; and (14) a method for producing the above polypeptide. NOVX sequences have antidiabetic, anorectic, antibacterial, virucide, immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian and antilipemic activities, and can be used in gene therapy. The polypeptide is useful in manufacturing a medicament for treating a syndrome associated with a human disease. The polypeptide or the nucleic acid molecule may be used to diagnose, treat or prevent metabolic disorders such as diabetes or obesity, infections, cachexia, cancer, neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease, immune disorders, haematopoietic disorders and various dyslipidaemias. The nucleic acids can also be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The present sequence encodes a human NOVX protein from the present invention.

Db 361 GACATCTATGCTATGAAAGTGATGAAGAAGGCTTTTATTGGCCACGAGCAGGTTTCA 420  
QY 421 TTTTGTGAGGAAGAGCGGCAACATATTTATCTCGAAGCACAAGCCGTGGATCCCCAAATTA 480  
Db 421 TTTTGTGAGGAAGAGCGGCAACATATTTATCTCGAAGCACAAGCCGTGGATCCCCAAATTA 480  
QY 481 CAGTATGCCCTTTCAGGACAAAAATCACCTTTATCTGATCGAGGAATATCAGCCTCGAGGG 540  
Db 481 CAGTATGCCCTTTCAGGACAAAAATCACCTTTATCTGATCGAATATCAGCCTCGAGGG 540  
QY 541 GACTTCTGTCTATTTTGAATAGATATGAGGACCAAGTTAGATGAATACTGATACAGTTT 600  
Db 541 GACTTCTGTCTATTTTGAATAGATATGAGGACCAAGTTAGATGAATACTGATACAGTTT 600  
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCACTGATGGGATACGTGATCGA 660  
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTTACAGCGTTTCACTGATGGGATACGTGATCGG 660  
QY 661 GACATCAAGCCTGAGAACATTCCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720  
Db 661 GACATCAAGCCTGAGAACATTCCTGTTGACCGCACAGGACACATCAAGCTGGTGGATTTT 720  
QY 721 GGNCTGCGCGGAAATGATTCAAACAGATGGTGAATGCCAACTCCCGATTGGGACC 780  
Db 721 GGNCTGCGCGGAAATGATTCAAACAA--GGTGAATGCCAACTCCCGATTGGGACC 777  
QY 781 CCAGATTACATGGCTCCTCGAAGTGTGACTGTGATGAACGGGATGGAAAGGACCTTAC 840  
Db 778 CCAGATTACATGGCTCCTCGAAGTGTGACTGTGATGAACGGGATGGAAAGGACCTTAC 837  
QY 841 GGCTCGAGCTGTGACTGTGGTGTGAGTGGGGGTGATGGCTATGAGATGATTTATGGGAGA 900  
Db 838 GGCTCGAGCTGTGACTGTGGTGTGAGTGGGGGTGATGGCTATGAGATGATTTATGGGAGA 897  
QY 901 TCCCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAATATGAAATTCACAGCG 960  
Db 898 TCCCCCTTCGACAGGGAACCTCTGCCAGAACCTTCAATAATATGAAATTCACAGCG 957  
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATCTGAATCAAAGC 1020  
Db 958 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTGACTTTCTTGATCTGATCTGAATCAAAGC 1017  
QY 1021 TTGTTGTGGGCGCAGAAAGAGACTGAAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1080  
Db 1018 TTGTTGTGGGCGCAGAAAGAGACTGAAAGTTTGAAGTCTTTGCTGCCATCCTTTCTTC 1077  
QY 1081 TCTAAAATTGACTGGAAACAACTTCGTAACTCTCTCCGCCCTTCGTTCCCACTCAAG 1140  
Db 1078 TCTAAAATTGACTGGAAACAACTTCGTAACTCTCTCCGCCCTTCGTTCCCACTCAAG 1137  
QY 1141 TCTGACGATGACCTCCAAATTTGATGAAACAGAGAAATTCGTGGGTTTATCCTCT 1200  
Db 1138 TCTGACGATGACCTCCAAATTTGATGAAACAGAGAAATTCGTGGGTTTATCCTCT 1197  
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTTTCG 1260  
Db 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGCTTTGTGGGTTTTCG 1257  
QY 1261 TACAGCAAGGCACTGGGGATTCCTGGTATGATCTGAGTCTGTTGTGCGGGTCTGACTCC 1320  
Db 1258 TACAGCAAGGCACTGGGGATTCCTGGTATGATCTGAGTCTGTTGTGCGGGTCTGACTCC 1317  
QY 1321 CCGTGCAGCTGAGCTCCATGGAAAGAAATCTTCATCAAAAGCAAGAGCTTACAAGAC 1380  
Db 1318 CCGTGCAGCTGAGCTCCATGGAAAGAAATCTTCATCAAAAGCAAGAGCTTACAAGAC 1377  
QY 1381 TCTCAGGACAAGTGTCAACAGATGGAGAGGAATGACCGGTTATACCGGAGGTGTCA 1440  
Db 1378 TCTCAGGACAAGTGTCAACAGATGGAGAGGAATGACCGGTTATACCGGAGGTGTCA 1437  
QY 1441 GAGTGTGAGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAGGCTTCAAGCTTCAAGCTCAGAG 1500  
Db 1438 GAGTGTGAGCTGTGCTTAGTCAAGAGAGGTGGAGCTGAGGCTTCAAGCTTCAAGCTCAGAG 1497

QY 1501 TCCCTCTGAGCAGGACCTTGTCTACCTCATACAGAAATGCAGTAGCTTAAAGCGAAGT 1560  
Db 1498 TCCCTCTGAGCAGGACCTTGTCTACCTCATACAGAAATGCAGTAGCTTAAAGCGAAGT 1557  
QY 1561 TTGAGCAAGCACGGATGGAGGTGTCCTCAGAGGATGACAAAGACTGCGAGCTTCTCAT 1620  
Db 1558 TTGAGCAAGCACGGATGGAGGTGTCCTCAGAGGATGACAAAGACTGCGAGCTTCTCAT 1617  
QY 1621 GATATCAGAGCAGGACCGGAACTCCAAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1680  
Db 1618 GATATCAGAGCAGGACCGGAACTCCAAGAAATCAAGAGCAGGAGTACCAGGCTCAA 1677  
QY 1681 GTGAAAGAAATAGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1740  
Db 1678 GTGAAAGAAATAGAGTTGATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAGA 1737  
QY 1741 CGGAGTATCTCTACGAATCTGAGCTGAGAGGTCCTCGGCTTGTCTGGAAGATTCAG 1800  
Db 1738 CGGAGTATCTCTACGAATCTGAGCTGAGAGGTCCTCGGCTTGTCTGGAAGATTCAG 1797  
QY 1801 CGGAAAGCGACAGAAATGTCAAGATAAACTGTGAAGCTTAAGGATCAAGGGAAGCCTGAA 1860  
Db 1798 CGGAAAGCGACAGAAATGTCAAGATAAACTGTGAAGCTTAAGGATCAAGGGAAGCCTGAA 1857  
QY 1861 GTGGAGAAATATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920  
Db 1858 GTGGAGAAATATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1917  
QY 1921 CTCGAGAGAAACTGGAGAGGCT----- 1944  
Db 1918 CTCGAGAGAAACTGGAGAGGCTGTAAAGCCAGCAGGAGCCACCGAGCTGCTGCAG 1977  
QY 1945 -----GAAAAGGAGCGAGCGAGGAGCTGGAGAACTGGAGAACCCGAGAG 1992  
Db 1978 AATATCCGCCAGGAAAAGGAGCGAGCGAGGAGGCTGGAGAACTGCGAGAACCCGAGAG 2037  
QY 1993 GATTCTTTGAGGCATCAGAAAGAGCTGTGGAAGCTGAGGAACCGCGCATTTCTCTG 2052  
Db 2038 GATTCTTTGAGGCATCAGAAAGAGCTGTGGAAGCTGAGGAACCGCGCATTTCTCTG 2097  
QY 2053 GAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAAACAGACTGAAGGATGAC 2112  
Db 2098 GAGAACAGGTAAAGAGACTAGAGACCATGGAGCGTAGAGAAAACAGACTGAAGGATGAC 2157  
QY 2113 ATCGAGCAAAATCCAAACAGATCCAGCAGATGCTGATATAATTTCTGGAGCTCGAAGAG 2172  
Db 2158 ATCCAGCAAAATCCAAACAGATCCAGCAGATGCTGATATAATTTCTGGAGCTCGAAGAG 2217  
QY 2173 AAAATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAACTGCACTGAAAACAGAAAGAG 2232  
Db 2218 AAAATCGGGAGGCCCAAGTCTCAGCCAGCACCTAGAACTGCACTGAAAACAGAAAGAG 2277  
QY 2233 CAGCACTATAGGAAAAGATTAAAGTGTGGACAATCAGATAAAGAAAGACCTGGCTGAC 2292  
Db 2278 CAGCACTATAGGAAAAGATTAAAGTGTGGACAATCAGATAAAGAAAGACCTGGCTGAC 2337  
QY 2293 AAGGAGACACTGGAGAACATGATCGAGACACGAGGAGGAGGCCCATGAGAAGGGCAAA 2352  
Db 2338 AAGGAGACACTGGAGAACATGATCGAGACACGAGGAGGAGGCCCATGAGAAGGGCAAA 2397  
QY 2353 ATTCTCAGCGAACAGAGGCGATCAATGCTATGATTCCAAAGATCAGATCCCTGAA 2412  
Db 2398 ATTCTCAGCGAACAGAGGCGATCAATGCTATGATTCCAAAGATCAGATCCCTGAA 2457  
QY 2413 CAGAGGATTTGGAACTGTCTGAGCCAAATAAATTCGCAAAATACAGCTTTTATACC 2472  
Db 2458 CAGAGGATTTGGAACTGTCTGAGCCAAATAAATTCGCAAAATACAGCTTTTATACC 2517  
QY 2473 CAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTAC 2532  
Db 2518 CAAGGAAACATGAAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAACAGAAATTTAC 2577

QY 2533 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2592  
Db |||||  
2578 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAAACCGAAAACTGGAGGAGCAGCTGGAG 2637  
QY |||||  
2593 AAGATCAGCCACCAAGACCAAGTGCAGAAATCGGCTGCTGGAACTGGAGACAAGATTG 2652  
Db |||||  
2638 AAGATCAGCCACCAAGACCAAGTGCAGAAATCGGCTGCTGGAACTGGAGACAAGATTG 2697  
QY |||||  
2653 CGGAGGTCAGTCTAGAGACGAGGAGCAGAAACTGGAGTCAAGGCCAGCTCACAGAG 2712  
Db |||||  
2698 CGGAGGTCAGTCTAGAGACGAGGAGCAGAAACTGGAGTCAAGGCCAGCTCACAGAG 2757  
QY |||||  
2713 CTACAGCTCTCCCTGCGAGGCGGAGTGCAGATTCAGACCTGCGAGCTGCAAGGCCG 2772  
Db |||||  
2758 CTACAGCTCTCCCTGCGAGGCGGAGTGCAGATTCAGACCTGCGAGCTGCAAGGCCG 2817  
QY |||||  
2773 GCCTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGTGAAGAGACCAAGCT 2832  
Db |||||  
2818 GCCTGAGAGCCAGCTTCGCCAGGCGAAGACAGAGTGAAGAGACCAAGCT 2877  
QY |||||  
2833 GAAGAGAGATCAGGACCTCAGGCACATAGAGATGAATCCAGCGCAAAATTGATGCT 2892  
Db |||||  
2878 GAAGAGAGATCAGGACCTCAGGCACATAGAGATGAATCCAGCGCAAAATTGATGCT 2937  
QY |||||  
2893 CTTGTAACAGCTGTACTGTAACTCAGACCTGGAGAGCAGCTAAACAGCTGACCGAG 2952  
Db |||||  
2938 CTTGTAACAGCTGTACTGTAACTCAGACCTGGAGAGCAGCTAAACAGCTGACCGAG 2997  
QY |||||  
2953 GACAACTGAACTCAACAAACCAAACTTCTATCTGTCAAACAACTGATGAGGCTTCT 3012  
Db |||||  
2998 GACAACTGAACTCAACAAACCAAACTTCTATCTGTCAAACAACTGATGAGGCTTCT 3057  
QY |||||  
3013 GCGCCAAACACAGATTGTACAACTGCGAAGTGAAGTGAACCTCTCCCGCGGAGATC 3072  
Db |||||  
3058 GCGCCAAACACAGATTGTACAACTGCGAAGTGAAGTGAACCTCTCCCGCGGAGATC 3117  
QY |||||  
3073 ACAGAACGAGAGATGCACTTACAGCCAGAAACCAACGATGGAGCTCTGAAGACCAAG 3132  
Db |||||  
3118 ACAGAACGAGAGATGCACTTACAGCCAGAAACCAACGATGGAGCTCTGAAGACCAAG 3177  
QY |||||  
3133 TGCACCATCTGAGAGACAGGTCATGGATTGGAGGCCCTAAACGATGAGCTGTAGAA 3192  
Db |||||  
3178 TGCACCATCTGAGAGACAGGTCATGGATTGGAGGCCCTAAACGATGAGCTGTAGAA 3237  
QY |||||  
3193 AAGAGCGGAGTGGAGGCTCGAGAGCGCTCCTGGGTGATGAGAAATCCCAAGTTTGAG 3252  
Db |||||  
3238 AAGAGCGGAGTGGAGGCTCGAGAGCGCTCCTGGGTGATGAGAAATCCCAAGTTTGAG 3297  
QY |||||  
3253 TGTGGGTTGAGAGCTGAGAGATGCTGGACACCGAGAAACAGAGAGGGGAGAGCC 3312  
Db |||||  
3298 TGTGGGTTGAGAGCTGAGAGATGCTGGACACCGAGAAACAGAGAGGGGAGAGCC 3357  
QY |||||  
3313 GATCAGCGGATCACCAGCTCCCGAGGTGGTGGAGCTGGAGTGAAGGAGCAGAGGCT 3372  
Db |||||  
3358 GATCAGCGGATCACCAGCTCCCGAGGTGGTGGAGCTGGAGTGAAGGAGCAGAGGCT 3417  
QY |||||  
3373 GAGATCTGCTCTGAGAGGCTCTCAAGAGCAGAACTGAAAGCCAGAGCCTTCTCT 3432  
Db |||||  
3418 GAGATCTGCTCTGAGAGGCTCTCAAGAGCAGAACTGAAAGCCAGAGCCTTCTCT 3477  
QY |||||  
3433 GACAAGCTCAATGACCTGAGAGAGAGATGCTATGCTTGAATGAATGATCCCGAGCTTA 3492  
Db |||||  
3478 GACAAGCTCAATGACCTGAGAGAGAGATGCTATGCTTGAATGAATGATCCCGAGCTTA 3537  
QY |||||  
3493 CAGCAGAGCTGAGACTGAACGAGAGCTCAACGAGGCTCTGGAAGAGCAAGCCAAA 3552  
Db |||||  
3538 CAGCAGAGCTGAGACTGAACGAGAGCTCAACGAGGCTCTGGAAGAGCAAGCCAAA 3597  
QY |||||  
3553 TTACAGCAGAGATGACCTGAGAGAAATCAATTTCCGCTCTGACTCAAGGATGCAA 3612  
Db |||||  
3598 TTACAGCAGAGATGACCTGAGAGAAATCAATTTCCGCTCTGACTCAAGGATGCAA 3657  
QY |||||  
3613 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGAGTATCAGCTG 3672

Db |||||  
3658 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGAGTATCAGCTG 3717  
QY |||||  
3673 GAAAAATTCTAGGTTCTCTATTCTCATGAAAGGTGAAAAATGGAAGGCACTATTCTCAA 3732  
Db |||||  
3718 GAAAAATTCTAGGTTCTCTATTCTCATGAAAGGTGAAAAATGGAAGGCACTATTCTCAA 3777  
QY |||||  
3733 CAAACCAAACTCATTTGATTTCTGCAAGCCAAAAATGGAACCTGCTTAAAAAGAAAAAG 3792  
Db |||||  
3778 CAAACCAAACTCATTTGATTTCTGCAAGCCAAAAATGGAACCTGCTTAAAAAGAAAA-- 3835  
QY |||||  
3793 GGTTTATTAGTCGACGGAAGAGACCTCTTTACCCACACAGGTTCTCTCCAGTAC 3852  
Db |||||  
3836 -----AGGTGCTCTCTGAGTAC 3852  
QY |||||  
3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTAGAGAAAGCC 3912  
Db |||||  
3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTAGAGAAAGCC 3912  
QY |||||  
3913 CTTGAGAGACCCGATCGAGCTCGGTCCGCCCGGAGGAGCTGCCACCGCAAGCA 3972  
Db |||||  
3913 CTTGAGAGACCCGATCGAGCTCGGTCCGCCCGGAGGAGCTGCCACCGCAAGCA 3972  
QY |||||  
3973 ACAGACCAACCAACCCATCCACGCCACACCGGAGGAGAGATCGCCATCTCCGCC 4032  
Db |||||  
3973 ACAGACCAACCAACCCATCCACGCCACACCGGAGGAGAGATCGCCATCTCCGCC 4032  
QY |||||  
4033 ATCTGCGGTGCGCAGAGACACAGCCAGTGCATGAGCTGTGGCCCGCCATCCAGC 4092  
Db |||||  
4033 ATCTGCGGTGCGCAGAGACACAGCCAGTGCATGAGCTGTGGCCCGCCATCCAGC 4092  
QY |||||  
4093 CGCAGAAAGGAGTCTTCAAATCCAGAGAAATTTAGTGGGTCTTAAAGAACGATGCAC 4152  
Db |||||  
4093 CGCAGAAAGGAGTCTTCAAATCCAGAGAAATTTAGTGGGTCTTAAAGAACGATGCAC 4152  
QY |||||  
4153 CACAATATCTCACCAGATTCACGTAGGACTGAACATGCGAGCCACAAAGTGTGTGTG 4212  
Db |||||  
4153 CACAATATCTCACCAGATTCACGTAGGACTGAACATGCGAGCCACAAAGTGTGTGTG 4212  
QY |||||  
4213 TGTCTGATACCGTGCACTTTGGAGCCAGGCATCCAAATGCTCTCGAATCTCAGGTTAG 4272  
Db |||||  
4213 TGTCTGATACCGTGCACTTTGGAGCCAGGCATCCAAATGCTCTAGAATCTCAGGTTAG 4272  
QY |||||  
4273 TGTCAACCCAGTCTCCAGTCTTCCAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4332  
Db |||||  
4273 TGTCAACCCAGTCTCCAGTCTTCCAGCCACCTGCGGCTTGCCTGCTGAATATGCC 4332  
QY |||||  
4333 ACACACTTCCAGGAGGCTTCCCGTGACAAATGAATCCCGAGGCTCCAGACCAAG 4392  
Db |||||  
4333 ACACACTTCCAGGAGGCTTCCCGTGACAAATGAATCCCGAGGCTCCAGACCAAG 4392  
QY |||||  
4393 GAGCCAGCAGCAGCTTGCACCTGGAAAGGTTGATGAAGGTGCCAGGAATAACAAACGA 4452  
Db |||||  
4393 GAGCCAGCAGCAGCTTGCACCTGGAAAGGTTGATGAAGGTGCCAGGAATAACAAACGA 4452  
QY |||||  
4453 GGAAGCAAGGCTGGGACAGAAATGATCATTTGCTGGAGGATCAAAAGTCTCTATTAT 4512  
Db |||||  
4453 GGAAGCAAGGCTGGGACAGAAATGATCATTTGCTGGAGGATCAAAAGTCTCTATTAT 4512  
QY |||||  
4513 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAAAGATTTGAGCTGTGCTTCCC 4572  
Db |||||  
4513 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGGTGGAAAGATTTGAGCTGTGCTTCCC 4572  
QY |||||  
4573 GAGGGGATGATCTATTCTATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4632  
Db |||||  
4573 GAGGGGATGATCTATTCTATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4632  
QY |||||  
4633 GCA----- 4635  
Db |||||  
4633 GCAGATGCTCCATACACTACTGAAGATGGAATCTCACCCGACACACCTGCTGCTGCTGCTGCT 4692  
QY |||||  
4636 ----- 4636

Db 4693 AGAACCTCTACTTGTCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTACACGCCCTTA 4752  
Qy 4636 -----GAAAAAGCAGAGCTGATGCTAAACTG 4662  
Db 4753 GAATCAGTTGTGCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAACTGATGCTAAACTG 4812  
Qy 4663 CTTGGAACCTCCCTGCTGATAACTGGAAGTGTATGACCGTCTAGACATGAACCTGACACGCTG 4722  
Db 4813 CTTGGAACCTCCCTGCTGTAACCTGGAAGTGTATGACCGTCTAGACATGAACCTGACACGCTG 4872  
Qy 4723 CCCTTCAGTGACCAAGTGTGTGTGGGCAACGAGGAAGGGCTCTACGCCCTGAATGTC 4782  
Db 4873 CCCTTCAGTGACCAAGTGTGTGTGGGCAACGAGGAAGGGCTCTACGCCCTGAATGTC 4932  
Qy 4783 TTGAAAACTCCCTAACCATGCTCCAGAAATTGAGCAGTCTTCCAAATTTATATTATC 4842  
Db 4933 TTGAAAACTCCCTAACCATGCTCCAGAAATTGAGCAGTCTTCCAAATTTATATTATC 4992  
Qy 4843 AAGGACCTCGAAGAGTACTCATGTATAGCAGGAGAGCGGCGACTGTGTCTTGTGGAC 4902  
Db 4993 AAGGACCTCGAAGAGTACTCATGTATAGCAGGAGAGCGGCGACTGTGTCTTGTGGAC 5052  
Qy 4903 GTGAGAAAGTGAACAGTCCCTGGCCGAGTCCACCTGCTGCCAGCGCGACATCTCA 4962  
Db 5053 GTGAGAAAGTGAACAGTCCCTGGCCGAGTCCACCTGCTGCCAGCGCGACATCTCA 5112  
Qy 4963 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTGGGCGCAGCAAGATTGAGAAC 5022  
Db 5113 CCCAACATTTTGAAGCTGTCAAGGCTGCCACTTGTGTGGGCGCAGCAAGATTGAGAAC 5172  
Qy 5023 GGGCTCTGCATCTGTGCAGCCATGCCCAGCAAGTGTCTATTCTCCGGCTACAAACGAAAC 5082  
Db 5173 GGGCTCTGCATCTGTGCAGCCATGCCCAGCAAGTGTCTATTCTCCGGCTACAAACGAAAC 5232  
Qy 5083 CTCAGCAATATCTGATCCGGAAGAGATAGAGCTCAGAGCTCAGAGCTGTGATGTCAC 5142  
Db 5233 CTCAGCAATATCTGATCCGGAAGAGATAGAGCTCAGAGCTGTGATGTCAC 5292  
Qy 5143 TTCACCAATTACAGTATCTCTATTGGAACCAATAATTCTACGAAATCAGCATGAAGCAG 5202  
Db 5293 TTCACCAATTACAGTATCTCTATTGGAACCAATAATTCTACGAAATCAGCATGAAGCAG 5352  
Qy 5203 TACAGCTTGAGGAATCTCTGATAGATGACCATCTCTGGCAGCTCTGTGTTGCC 5262  
Db 5353 TACAGCTTGAGGAATCTCTGATAGATGACCATCTCTGGCAGCTCTGTGTTGCC 5412  
Qy 5263 GCCTCTTCAACAGCTTCCCTGCTCAATCGTGCAGTGAACAGCGCAGCGGAGAGAG 5322  
Db 5413 GCCTCTTCAACAGCTTCCCTGCTCAATCGTGCAGTGAACAGCGCAGCGGAGAG 5472  
Qy 5323 GAGTACTGTGTGTTTCCACGAATTGGAGTGTTCGTGGAATTTTACGGAGAGAGTAGC 5382  
Db 5473 GAGTACTGTGTGTTTCCACGAATTGGAGTGTTCGTGGAATTTTACGGAGAGAGTAGC 5532  
Qy 5383 CGCAGACGATCTCAAGTGGTGTCTTACCTTTGCGCTTTGCGCTTACAGAGAACCTAT 5442  
Db 5533 CGCAGACGATCTCAAGTGGTGTCTTACCTTTGCGCTTTGCGCTTACAGAGAACCTAT 5592  
Qy 5443 CTGTTTGTGACCCCACTTCAACTCACTCGAAGTAATTGAGATCAGGCACGCTCTCAGCA 5502  
Db 5593 CTGTTTGTGACCCCACTTCAACTCACTCGAAGTAATTGAGATCAGGCACGCTCTCAGCA 5652  
Qy 5503 GGGACCCCTCCGAGGTACCTGGACATCCGAAACCGCGCTTACCTGGGCGCTGCAATT 5562  
Db 5653 GGGACCCCTCCGAGGTACCTGGACATCCGAAACCGCGCTTACCTGGGCGCTGCAATT 5712  
Qy 5563 TCCTCAGGAGGATTTACTTGGGCTCCTCATACAGGATAAATTAAAGGGTCAATTGCTGC 5622  
Db 5713 TCCTCAGGAGGATTTACTTGGGCTCCTCATACAGGATAAATTAAAGGGTCAATTGCTGC 5772  
Qy 5623 AAGGAAACCTGTGGAAGAGTCCGGCACTGTAACAACCGGGGCGCTCCACCTCCCGG 5682  
Db 5773 AAGGAAACCTGTGGAAGAGTCCGGCACTGTAACAACCGGGGCGCTCCACCTCCCGG 5832

Qy 5683 AGCAGCCCCAACAGCGAGGCCACCCACGTAACAACGAGCACATCACCAAGCGCTGGCC 5742  
Db 5833 AGCAGCCCCAACAGCGAGGCCACCCACGTAACAACGAGCACATCACCAAGCGCTGGCC 5892  
Qy 5743 TCCAGCCCCAGCGCGCGCGAAGCGCCAGCCACCCGCGAGAGCCAGCACACCCACCGC 5802  
Db 5893 TCCAGCCCCAGCGCGCGCGAAGCGCCAGCCACCCGCGAGAGCCAGCACACCCACCGC 5952  
Qy 5803 TACCAGCGGGCGGACCCGAGCTGCGCAGGACAAGTCTCTGCGCGCCCTTGGAGCGA 5862  
Db 5953 TACCAGCGGGCGGACCCGAGCTGCGCAGGACAAGTCTCTGCGCGCCCTTGGAGCGA 6012  
Qy 5863 GAGAAGTCCCGCGCGCGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGCGCTTTT 5922  
Db 6013 GAGAAGTCCCGCGCGCGATGCTCAGCACGCGGAGAGCGGTCCCGCGGAGCGCTTTT 6072  
Qy 5923 GAAGACAGCAGCAGGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGCCAGGTG 5982  
Db 6073 GAAGACAGCAGCAGGGCGCGCTGCTGCGGAGCGCTGAGGACCCCGCTGCCAGGTG 6132  
Qy 5983 AACAAAGGAAGAGGGCA 5999  
Db 6133 AACAAAGTGGAGCGCA 6149

## RESULT 12

ABQ78871  
ID ABQ78871 standard; cDNA; 5877 BP.  
XX AC ABQ78871;  
XX DT 10-OCT-2002 (first entry)  
XX DE Human kinase cDNA #2.  
XX KW Human; kinase; enzyme; serine-threonine kinase; nootropic; cytostatic;  
KW Citron rho-interacting kinase; gene therapy; mental disorder; cancer;  
KW gene; ss.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT CDS 1..5877  
FT /\*tag= a  
FT /product= "Kinase"  
XX WO200259325-A2.  
XX PN 01-AUG-2002.  
XX PD 20-DEC-2001; 2001WO-US050497.  
XX PF 27-DEC-2000; 2000US-0258335P.  
XX PR (LEXI-) LEXICON GENETICS INC.  
XX PA Yu X, Miranda M, Fridele CJ;  
XX PI WPI; 2002-599796/64.  
XX DR P-PSDB; ABB81928.  
XX DR Novel polynucleotide encoding human proteins that are structurally  
XX similar to animal kinases, useful for drug screening, diagnosis, in gene  
XX therapy of disorders and diseases e.g. cancer and pharmacogenomic  
XX applications.  
XX PS Disclosure; Page 44-45; 50pp; English.  
XX CC The invention relates to a novel human protein that shares structural  
XX similarity with animal kinases, including serine-threonine kinases,  
XX particularly Citron rho-interacting kinases. The proteins of the  
XX invention have nootropic and cytostatic activity. The polynucleotides may

CC have a use in gene therapy. The encoded novel polypeptides are useful for  
 CC generating antibodies, as reagents in diagnostic assays, for identifying  
 CC other cellular gene products related to NHP and as reagents in assays for  
 CC screening for compounds that are useful in the treatment of mental,  
 CC biological or medical disorders and diseases including cancer. The  
 CC sequence encodes a novel human kinase of the invention  
 XX  
 SQ

Sequence 5877 BP; 1680 A; 1433 C; 1583 G; 1181 T; 0 U; 0 Other;

Query Match 87.0%; Score 5358; DB 6; Length 5877;  
 Best Local Similarity 95.8%; Pred. No. 0;  
 Matches 5634; Conservative 0; Mismatches 5; Indels 243; Gaps 3;

QY 1 ATGTTGAAGTCAATATGAGCGCGGAATCCCTTGGATGCTGGTCTGTAACCCATT 60  
 DB |||||  
 QY 1 ATGTTGAAGTCAATATGAGCGCGGAATCCCTTGGATGCTGGTCTGTAACCCATT 60  
 DB |||||  
 QY 61 GCCAGCGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCCCTTATGACTCAA 120  
 DB |||||  
 QY 61 GCCAGCGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCCCTTATGACTCAA 120  
 DB |||||  
 QY 121 CAGCAGATGCTCTCTTCCGAGAGGAGATATAGATGCCCTCTTTGTTCTTTTGA 180  
 DB |||||  
 QY 121 CAGCAGATGCTCTCTTCCGAGAGGAGATATAGATGCCCTCTTTGTTCTTTTGA 180  
 DB |||||  
 QY 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGGAAGTAT 240  
 DB |||||  
 QY 181 GAATGCACTCAGCTGCTCTGATGAAGATTAAAGCAGCTGAGCAACTTTGTCGGGAAGTAT 240  
 DB |||||  
 QY 241 TCGCAGACCATAGCTGAGTTACAGGAGCTCCAGCTTGGCAGAAAGACTTCGAAGTCAGA 300  
 DB |||||  
 QY 241 TCGCAGACCATAGCTGAGTTACAGGAGCTCCAGCTTGGCAGAAAGACTTCGAAGTCAGA 300  
 DB |||||  
 QY 301 AGTCTTGAGTGTGCTGCTCTGATGAAGTTAAAGCAGCTGAGCAACTTTGTCGGGAAGTAT 360  
 DB |||||  
 QY 301 AGTCTTGAGTGTGCTGCTCTGATGAAGTTAAAGCAGCTGAGCAACTTTGTCGGGAAGTAT 360  
 DB |||||  
 QY 361 GACATCTATGCTATGAAGTGTGAAGAGAGGCTTTATGGCCAGGAGAGGTTTCA 420  
 DB |||||  
 QY 361 GACATCTATGCTATGAAGTGTGAAGAGAGGCTTTATGGCCAGGAGAGGTTTCA 420  
 DB |||||  
 QY 421 TTTTGTGAGAGAGCGGAACATATTTATCTGAGAGCAGAGCCGCTGGATCCGCCAATTA 480  
 DB |||||  
 QY 421 TTTTGTGAGAGAGCGGAACATATTTATCTGAGAGCAGAGCCGCTGGATCCGCCAATTA 480  
 DB |||||  
 QY 481 CAGTATGCTCTTCAGGACAAATAATCACTTTATCTGATGGAGAAATATCAGCTCGAGGG 540  
 DB |||||  
 QY 481 CAGTATGCTCTTCAGGACAAATAATCACTTTATCTGATGGAGAAATATCAGCTCGAGGG 540  
 DB |||||  
 QY 541 GACTTGTGCTCACTTTTGAATAGATAGAGGACAGTTAGATGAATACTGATACAGTTT 600  
 DB |||||  
 QY 541 GACTTGTGCTCACTTTTGAATAGATAGAGGACAGTTAGATGAATACTGATACAGTTT 600  
 DB |||||  
 QY 601 TACTAGCTGAGCTGATTTTGGCTGTTTACAGGTTTCTATCTGATGGATACGTCATCGA 660  
 DB |||||  
 QY 601 TACTAGCTGAGCTGATTTTGGCTGTTTACAGGTTTCTATCTGATGGATACGTCATCGA 660  
 DB |||||  
 QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGATTT 720  
 DB |||||  
 QY 661 GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGGTGATTT 720  
 DB |||||  
 QY 721 GGAATCTGCCGGAATGAATCAACAGATGGTGAATGCGCAACTCCGATTTGGGACC 780  
 DB |||||  
 QY 721 GGAATCTGCCGGAATGAATCAACAGATGGTGAATGCGCAACTCCGATTTGGGACC 780  
 DB |||||  
 QY 781 CCAGATTACATGGCTCTGAGTGTGATGATGACGCGGATCGAAAGGACCTTAC 840  
 DB |||||  
 QY 781 CCAGATTACATGGCTCTGAGTGTGATGATGACGCGGATCGAAAGGACCTTAC 840  
 DB |||||  
 QY 841 GGCCTGGACTGTGCTGCTGAGTGGCTGATGCTATGAGATGATTTATGGGAGA 900  
 DB |||||  
 QY 841 GGCCTGGACTGTGCTGCTGAGTGGCTGATGCTATGAGATGATTTATGGGAGA 900  
 DB |||||

QY 901 TCCCCCTTGCAGAGGAAACCTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 960  
 DB |||||  
 QY 901 TCCCCCTTGCAGAGGAAACCTCTGCCAGAACCTTCAATAACATTATGAATTTCCAGCGG 960  
 DB |||||  
 QY 961 TTTTGAATTTCCAGATGACCCCAAGTGGAGTGAATTTCTTGATCTGATTTCAAGC 1020  
 DB |||||  
 QY 961 TTTTGAATTTCCAGATGACCCCAAGTGGAGTGAATTTCTTGATCTGATTTCAAGC 1020  
 DB |||||  
 QY 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGTCGCATCTCTTTCTTC 1080  
 DB |||||  
 QY 1021 TTGTTGTGCGGCCAGAAAGAGAGACTGAAGTTTGAAGTCTTTGTCGCATCTCTTTCTTC 1080  
 DB |||||  
 QY 1081 TCTAAATTTGATGAGAACCAATCTGTAACCTCTCTCCCTCTGTTTCCACCTCTCAAG 1140  
 DB |||||  
 QY 1081 TCTAAATTTGATGAGAACCAATCTGTAACCTCTCTCCCTCTGTTTCCACCTCTCAAG 1140  
 DB |||||  
 QY 1141 TCTGACGATGACCTCCAAATTTTGTGATGAACAGAGAGAAATTCGTGGTTCATCTCT 1200  
 DB |||||  
 QY 1141 TCTGACGATGACCTCCAAATTTTGTGATGAACAGAGAGAAATTCGTGGTTCATCTCT 1200  
 DB |||||  
 QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAGAACTGCGGTTTGTGGGTTTTCG 1260  
 DB |||||  
 QY 1201 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAGAACTGCGGTTTGTGGGTTTTCG 1260  
 DB |||||  
 QY 1261 TACAGAGGCACTGGGATTTCTGTTAGATCTGAGTCTGTTGTCGGTCTCGACTCC 1320  
 DB |||||  
 QY 1261 TACAGAGGCACTGGGATTTCTGTTAGATCTGAGTCTGTTGTCGGTCTCGACTCC 1320  
 DB |||||  
 QY 1321 CCTGCCAAGCTAGCTCCATGGAAGAAATCTTCTCAATAAAGCAAGAGCTACAAGAC 1380  
 DB |||||  
 QY 1321 CCTGCCAAGCTAGCTCCATGGAAGAAATCTTCTCAATAAAGCAAGAGCTACAAGAC 1380  
 DB |||||  
 QY 1381 TCTCAGGCAAGTGTCAAGATGAGCAGAAATGACCGGTTACATCGGAGAGTGTCA 1440  
 DB |||||  
 QY 1381 TCTCAGGCAAGTGTCAAGATGAGCAGAAATGACCGGTTACATCGGAGAGTGTCA 1440  
 DB |||||  
 QY 1441 GAGTGGAGGCTGTCTTAGTCAAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
 DB |||||  
 QY 1441 GAGTGGAGGCTGTCTTAGTCAAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500  
 DB |||||  
 QY 1501 TCCCTCTCGAGCAGGACCTTGTACCTACATACAGAAATGAGTAGCTTAAAGCGAGT 1560  
 DB |||||  
 QY 1501 TCCCTCTCGAGCAGGACCTTGTACCTACATACAGAAATGAGTAGCTTAAAGCGAGT 1560  
 DB |||||  
 QY 1561 TTGAGCAAGCAGGATGGAGGTGTCCAGAGAGATGACAAAGCAGCTGAGGCTTCCAT 1620  
 DB |||||  
 QY 1561 TTGAGCAAGCAGGATGGAGGTGTCCAGAGAGATGACAAAGCAGCTGAGGCTTCCAT 1620  
 DB |||||  
 QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680  
 DB |||||  
 QY 1621 GATATCAGAGCAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGGAGTACAGGCTCAA 1680  
 DB |||||  
 QY 1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
 DB |||||  
 QY 1681 GTGGAAGAAATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
 DB |||||  
 QY 1741 CGGAGTGTCTACGAAATCTGAGTGTGAGAGTCTCGGCTTCTGCTGAGAAATCAAG 1800  
 DB |||||  
 QY 1741 CGGAGTGTCTACGAAATCTGAGTGTGAGAGTCTCGGCTTCTGCTGAGAAATCAAG 1800  
 DB |||||  
 QY 1801 CGGAAAGCAGAGATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1860  
 DB |||||  
 QY 1801 CGGAAAGCAGAGATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1860  
 DB |||||  
 QY 1861 GTGGGAGAAATATCGAAACTGGGAGAGATCAATGCTGAGCAGAGCTCAAAATTTCAAG 1920  
 DB |||||  
 QY 1861 GTGGGAGAAATATCGAAACTGGGAGAGATCAATGCTGAGCAGAGCTCAAAATTTCAAG 1920  
 DB |||||  
 QY 1921 CTCCAGAGAAATCTGAGAAAGGCT----- 1944  
 DB |||||  
 QY 1921 CTCCAGAGAAATCTGAGAAAGGCTGTAAAGCCAGCAGAGGACCCAGAGCTGTGTCAG 1980  
 DB |||||  
 QY 1945 -----GCAAAAGGAGCGAGCGGAGCTGGAGAGCTGGAGAAAGCTGAGAACCGAGAG 1992  
 DB |||||

Db 1981 AATATCCGCCGCAAGAGCGAGCCGAGAGGAGCTGGAGAGCTCGAAGACCGAG 2040  
Qy 1993 GATTTCTTCTGAAGGCATCAGAAAGAGCTGTGGAAGCTGAGGAACGCCGCATTTCTG 2052  
Db 2041 GATTTCTTCTGAAGGCATCAGAAAGAGCTGTGGAAGCTGAGGAACGCCGCATTTCTG 2100  
Qy 2053 GAGAACAGGTAAAGAGACTAGAGACCATGAGACCTGAGACCTAGAGAAACAGAGATGAC 2112  
Db 2101 GAGAACAGGTAAAGAGACTAGAGACCATGAGACCTGAGACCTAGAGAAACAGAGATGAC 2160  
Qy 2113 ATCCAGACAAAATCCCAACAGATCCAGCAGATGAGCTGATAAAATTTCTGGAGCTGAAAG 2172  
Db 2161 ATCCAGACAAAATCCCAACAGATCCAGCAGATGAGCTGATAAAATTTCTGGAGCTGAAAG 2220  
Qy 2173 AAAATCGGGAGGCCCAAGTCTCAGCCGAGCAGCTAGAGTCCACCTGAAACAGAAAGAG 2232  
Db 2221 AAAATCGGGAGGCCCAAGTCTCAGCCGAGCAGCTAGAGTCCACCTGAAACAGAAAGAG 2280  
Qy 2233 CAGCACTATGAGAAAAGATTAAAGTTGTTGGAATCAGATAAAGAAAGACCTGCTGAC 2292  
Db 2281 CAGCACTATGAGAAAAGATTAAAGTTGTTGGAATCAGATAAAGAAAGACCTGCTGAC 2340  
Qy 2293 AAGGAGACTCGAGAACATGATGAGAGACACGAGGAGGAGGCCATGAGAAAGGGCAAA 2352  
Db 2341 AAGGAGACTCGAGAACATGATGAGAGACACGAGGAGGAGGCCATGAGAAAGGGCAAA 2400  
Qy 2353 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGATGATCCAGATCAGATCCCTGGAA 2412  
Db 2401 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGATGATCCAGATCAGATCCCTGGAA 2460  
Qy 2413 CAGAGATTGTGGAACTGTCTGAAGCCAAATAAATTTGAGCAAAATAGCAGTCTTTTACC 2472  
Db 2461 CAGAGATTGTGGAACTGTCTGAAGCCAAATAAATTTGAGCAAAATAGCAGTCTTTTACC 2520  
Qy 2473 CAAAGGAAATGAGGCCCAAGAGAGATGATTTCTGAATCAGGCAACAGAAATTTTAC 2532  
Db 2521 CAAAGGAAATGAGGCCCAAGAGAGATGATTTCTGAATCAGGCAACAGAAATTTTAC 2580  
Qy 2533 CTGGAGACACAGGCTGGGAAGTTGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG 2592  
Db 2581 CTGGAGACACAGGCTGGGAAGTTGAGGCCAGAAACCGAAACTGGAGGAGCAGCTGGAG 2640  
Qy 2593 AAGATCAGCCACCAAGACCAAGTGAACAGATTCGGCTCTCGAATCTGGAGCAAGATTG 2652  
Db 2641 AAGATCAGCCACCAAGACCAAGTGAACAGATTCGGCTCTCGAATCTGGAGCAAGATTG 2700  
Qy 2653 CGGGAGTCACTAGAGCACGAGGAGAGAACTGGAGCTCAAGCGCCAGCTCACAGAG 2712  
Db 2701 CGGGAGTCACTAGAGCACGAGGAGAGAACTGGAGCTCAAGCGCCAGCTCACAGAG 2760  
Qy 2713 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCCCTGAGGCTGACCGGGCG 2772  
Db 2761 CTACAGCTCTCCCTGAGGAGCGGAGTCAAGTTGACAGCCCTGAGGCTGACCGGGCG 2820  
Qy 2773 GCCTGGAGAGCCAGCTTGCAGGCGAAGACAGAGCTGGAAGACCAACAGAGAGCT 2832  
Db 2821 GCCTGGAGAGCCAGCTTGCAGGCGAAGACAGAGCTGGAAGACCAACAGAGAGCT 2880  
Qy 2833 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2892  
Db 2881 GAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAAATCCAGCGCAAAATTTGATGCT 2940  
Qy 2893 CTTTCGTAACAGTGTACTGTAAATCAAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 2952  
Db 2941 CTTTCGTAACAGTGTACTGTAAATCAAGACCTGGAGGAGCAGCTAAACAGCTGACCGAG 3000  
Qy 2953 GACAACGCTGAACTCAACACCAAAATTTCTGTTCCAAACAACTCGATGAGGCTTCT 3012  
Db 3001 GACAACGCTGAACTCAACACCAAAATTTCTGTTCCAAACAACTCGATGAGGCTTCT 3060  
Qy 3013 GCGCCAAACGAGAGATTGTACAACTGCGAAGTGAAGTGGACCACTCCCGCCGGAGATC 3072

Db 3061 GCGCCAAACGAGAGATTGTACAACTGCGAAGTGAAGTGGACCACTCTCCCGCCGGAGATC 3120  
Qy 3073 ACGGAACGAGAGATGAGCTTACAGCCAGAAACGAAACGATGAGGCTCTGAGAACCAAG 3132  
Db 3121 ACGGAACGAGAGATGAGCTTACAGCCAGAAACGAAACGATGAGGCTCTGAGAACCAAG 3180  
Qy 3133 TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTAAACCGATGAGCTGCTAGAA 3192  
Db 3181 TGCACCATGCTGGAGGAACAGGTCATGGATTTGGAGGCCCTAAACCGATGAGCTGCTAGAA 3240  
Qy 3193 AAGAGCGGAGTGGAGGAGCTGAGAGAGCTGCTGGTGTATGAGAAATCCAGTTTGGAG 3252  
Db 3241 AAGAGCGGAGTGGAGGAGCTGAGAGAGCTGCTGGTGTATGAGAAATCCCAAGTTTGGAG 3300  
Qy 3253 TGTCCGGTTCGAGAGCTGAGAGAAATGCTCGACACCGAGAAACAGAGACGGGCGAGAGCC 3312  
Db 3301 TGTCCGGTTCGAGAGCTGAGAGAAATGCTCGACACCGAGAAACAGAGACGGGCGAGAGCC 3360  
Qy 3313 GATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAGAGGACCAAGGCT 3372  
Db 3361 GATCAGCGGATCACCGAGTCTCGCCAGGTGGTGGAGCTGGCAGTGAAGAGGACCAAGGCT 3420  
Qy 3373 GAGATTTCTGCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAGAGCCGAGAGCTCTCT 3432  
Db 3421 GAGATTTCTGCTCTGAGCAGGCTCTCAAGAGCAGAGCTGAGAGCCGAGAGCTCTCT 3480  
Qy 3433 GACAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAAATGAAATGCGCGAGCTTA 3492  
Db 3481 GACAAGCTCAATGACCTGGAGAAAGCATGCTATGCTTGAATGAAATGAAATGCGCGAGCTTA 3540  
Qy 3493 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGCAAA 3552  
Db 3541 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAGCAAA 3600  
Qy 3553 TTACAGCAGCAGATGAGACCTGCAAGAAATCAACATTTTCCGTCTGACTCAAGGACTCAA 3612  
Db 3601 TTACAGCAGCAGATGAGACCTGCAAGAAATCAACATTTTCCGTCTGACTCAAGGACTCAA 3660  
Qy 3613 GAAGCTTAGATCGGGTGTATCTACTGAAGACAGAAAGATGACTTGGAGATATCAGCTG 3672  
Db 3661 GAAGCTTAGATCGGGTGTATCTACTGAAGACAGAAAGATGACTTGGAGATATCAGCTG 3720  
Qy 3673 GAAGCATTGAGGTTCTTCTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3732  
Db 3721 GAAGCATTGAGGTTCTTCTATTTCTCATGAAAGGTGAAATGGAAGGCACTATTTCTCAA 3780  
Qy 3733 CAAACCAAACTCATTGATTTTCTGCAAGCCAAATGACCAACCTGCTTAAAGAGAAAG 3792  
Db 3781 CAAACCAAACTCATTGATTTTCTGCAAGCCAAATGACCAACCTGCTTAAAGAGAAAG -- 3838  
Qy 3793 GGTTTATTTAGTCAGCGAAAGGAGCCCTGCTTTACCCACACAGGTTCTCTGCGATAC 3852  
Db 3839 -----AGGTTCTCTGCGATAC 3855  
Qy 3853 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTAGAGGAGAGCC 3912  
Db 3856 AATGAGCTGAAGCTGGCCCTGGAGAGGAGAAAGCTCGCTGTGAGAGCTAGAGGAGAGCC 3915  
Qy 3913 CTTTCAAGAGACCCGATCGAGCTCCGCTCCGCGGAGGAGAGCTGCCACCGCAAGAGCA 3972  
Db 3916 CTTTCAAGAGACCCGATCGAGCTCCGCTCCGCGGAGGAGAGCTGCCACCGCAAGAGCA 3975  
Qy 3973 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGAGAGCTGCCATGCTCCGCC 4032  
Db 3976 ACGGACCAACCAACCCATCCAGCCAGCCACCGGAGGAGAGAGCTGCCATGCTCCGCC 4035  
Qy 4033 ATCGTGGGCTGGCCAGAGCAGCCAGCTGCCATGAGCTGTGTCGCTGCGCCGCTCAGC 4092  
Db 4036 ATCGTGGGCTGGCCAGAGCAGCCAGCTGCCATGAGCTGTGTCGCTGCGCCGCTCAGC 4095  
Qy 4093 CGCAGAAAGAGTCTTCACTCCAGAGAAATTTAGTGGGCTCTTAGGAGAGCAGATCCAC 4152  
Db 4096 CGCAGAAAGAGTCTTCACTCCAGAGAAATTTAGTGGGCTCTTAGGAGAGCAGATCCAC 4155

```
QY 4153 CACAAATATCTCTCACCGATTCAACGTAGGACTGAAATCATGGAGCCACAAAGTGTGCTGTG 4212
Db 4156 CACAATATCTCTCACCGATTCAACGTAGGACTGAAATCATGGAGCCACAAAGTGTGCTGTG 4215
QY 4213 TGTCTGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCCAGGTGATG 4272
Db 4216 TGTCTGATACCGTGCACCTTTGGACGCCAGGCATCCAAATGTCTCGAATGTCCAGGTGATG 4275
QY 4273 TGTACACCCCAAGTGTCTCCAGTCTTGCACAGCCACCTGCGGCTGTGCTGTAATATGCC 4332
Db 4276 TGTACACCCCAAGTGTCTCCAGTCTTGCACAGCCACCTGCGGCTGTGCTGTAATATGCC 4335
QY 4333 ACACATTTACCCGAGGCTTCTCCGCTGACAAATGAATCCCCAGGTCTCCAGACCAAG 4392
Db 4336 ACACATTTACCCGAGGCTTCTCCGCTGACAAATGAATCCCCAGGTCTCCAGACCAAG 4395
QY 4393 GAGCCACGACGAGCTTGCACTTGGAAAGGTGGATGAAGTGCACGGAATACCAACGA 4452
Db 4396 GAGCCACGACGAGCTTGCACTTGGAAAGGTGGATGAAGTGCACGGAATACCAACGA 4455
QY 4453 GGACAGCAAGCTGGACAGAGTACATTTGCTCGAGGGATCAAAAGTCTCATTTAT 4512
Db 4456 GGACAGCAAGCTGGACAGAGTACATTTGCTCGAGGGATCAAAAGTCTCATTTAT 4515
QY 4513 GACAATGAAGCCAGAGAGCTGGACAGAGCCCGTGGAAAGATTTGAGCTGTGCCCTTCC 4572
Db 4516 GACAATGAAGCCAGAGAGCTGGACAGAGCCCGTGGAAAGATTTGAGCTGTGCCCTTCC 4575
QY 4573 GACGGGATGTATCTATTCATTTGCTGCGCTTGTGCTTCCGAATTCGCAATACAGCCAAA 4632
Db 4576 GACGGGATGTATCTATTCATTTGCTGCGCTTGTGCTTCCGAATTCGCAATACAGCCAAA 4635
QY 4633 GCA----- 4635
Db 4636 GCAGATGTCCATACATCTGAAGATGGAATCTCACCCGACACACACTGCTGGCCCGGG 4695
QY 4636 ----- 4635
Db 4696 AGAACCTCTACTTGTAGTCTCCAGCTTCCCTGTGACAAACAGCGTGGGTCAACCGCTTA 4755
QY 4636 -----GAAAGACAGAGCTGATCTAAACTG----- 4662
Db 4756 GAATCAGTTGTCGACGTTGGAGAGTTTCTAGGAAAAAGCAGAGCTGATCTAAACTG 4815
QY 4663 CTTGAAACTCCTCTGCTGAACTGGAAGTGTATGACCGCTCTAGACATGAATGCACGCTG 4722
Db 4816 CTTGAAACTCCTCTGCTGAACTGGAAGTGTATGACCGCTCTAGACATGAATGCACGCTG 4875
QY 4723 CCCTTCAGTACACAGTGTGTGTTGGACCGAGGAGGCTCTACGCTGATGTC 4782
Db 4876 CCCTTCAGTACACAGTGTGTGTTGGACCGAGGAGGCTCTACGCTGATGTC 4935
QY 4783 TTGAAAACTCCCTAACCCATGTCCAGGAATGGAGCAGTCTTCCAAATTTATATTATC 4842
Db 4936 TTGAAAACTCCCTAACCCATGTCCAGGAATGGAGCAGTCTTCCAAATTTATATTATC 4995
QY 4843 AAGGACCTGGAGAGTACTCATGTATAGCAGAGAGAGCGGCACTGTGCTTGTGAC 4902
Db 4996 AAGGACCTGGAGAGTACTCATGTATAGCAGAGAGAGCGGCACTGTGCTTGTGAC 5055
QY 4903 GTCAAAAAAGTGAACAGTCCCTGGCCAGTCCCACTTCCCTGCGCCAGCCGACATCTCA 4962
Db 5056 GTGAGAAAGTGAACAGTCCCTGGCCAGTCCCACTTCCCTGCGCCAGCCGACATCTCA 5115
QY 4963 CCCAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5022
Db 5116 CCCAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTTGGGCGAGGCAAGATTGAGAAC 5175
QY 5023 GGGCTCTGATCTGTGACGCAATGCCAGCAAGTCTGCTTCTCCGCTCAACAGAAAC 5082
Db 5176 GGGCTCTGATCTGTGACGCAATGCCAGCAAGTCTGCTTCTCCGCTCAACAGAAAC 5235
```

```
QY 5083 CTCAGCAATTAAGTCTGATCCGAAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5142
Db 5236 CTCAGCAATTAAGTCTGATCCGAAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5295
QY 5143 TTCACCAATTAAGTCTGATCCGAAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5202
Db 5296 TTCACCAATTAAGTCTGATCCGAAAGAGATAGACCTCAGAGCCCTGAGCTGTATCCAC 5355
QY 5203 TACACGCTCGAGGAATTCCTGGATGAAGATGACCAATTCCTTTGGACACCTGTGTGTTGCC 5262
Db 5356 TACACGCTCGAGGAATTCCTGGATGAAGATGACCAATTCCTTTGGACACCTGTGTGTTGCC 5415
QY 5263 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCTGAGGTGAACAGCGCAGGCGCAGAG 5322
Db 5416 GCCTCTTCCAAACAGCTTCCCTGTCTCAATCTGAGGTGAACAGCGCAGGCGCAGAG 5475
QY 5323 GAGTACTTGTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5382
Db 5476 GAGTACTTGTGCTGTGTTTCCACGAATTTGGAGTGTTCGTGATTTCTTACGGAAGACGTAGC 5535
QY 5383 CGCAGACGATCTCAAGTGGAGTCCCTTACCTTTGGCCTTTGCTTACAGAAACCTAT 5442
Db 5536 CGCAGACGATCTCAAGTGGAGTCCCTTACCTTTGGCCTTTGCTTACAGAAACCTAT 5595
QY 5443 CTGTTTGTGACCCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGCACGCTCCTCAGCA 5502
Db 5596 CTGTTTGTGACCCCACTTCAACTCACTCGAAGTAATTGAGATCCAGGCACGCTCCTCAGCA 5655
QY 5503 GGGACCCCTGCCGAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCAATT 5562
Db 5656 GGGACCCCTGCCGAGCGTACCTGGACATCCGAAACCCGCGCTACCTGGGCGCTGCAATT 5715
QY 5563 TCCTCAGAGCGATTTACTTTGGGCTCTCTATACAGGATAAATAAGGGTCAATTGCTGC 5622
Db 5716 TCCTCAGAGCGATTTACTTTGGGCTCTCTATACAGGATAAATAAGGGTCAATTGCTGC 5775
QY 5623 AAGGAAACCTCGTGAAGAGTCCGGCACTGAACACACCGGGGCGCTGCCACCTCCGC 5682
Db 5776 AAGGAAACCTCGTGAAGAGTCCGGCACTGAACACACCGGGGCGCTGCCACCTCCGC 5835
QY 5683 AG 5684
Db 5836 AG 5837

RESULT 13
AAC77568
ID AAC77568 standard; cDNA; 6609 BP.
XX
AC AAC77568;
XX
DT 08-FEB-2001 (first entry)
XX
Human OREF3123 polynucleotide sequence SEQ ID NO:6245.
XX
Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
vulnerable; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
anticongestant; osteopathic; antiarthritic; immunosuppressant; cardiant;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antineoplastic;
antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
antianemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
FN WO2005473-A2.
```







1636 AGCCGAGCTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAAGTGAAGAAATGAGG 1695  
1936 AGCCGAGCTCCAGAAATCAAGAGCAGGAGTACCAGGCTCAAGTGAAGAAATGAGG 1877  
1696 TTGATGATGAATCAGTTGGAAGAGGATCTTTGTCTCAGCAAGAGAGCGAGTGTCTCTAC 1755  
1876 TTGATGATGAATCAGTTGGAAGAGGATCTTTGTCTCAGCAAGAGAGCGAGTGTCTCTAC 1817  
1756 GAATCTGAGCTGAGAGATCTCGGCTTGTCTGCTGAAGAAATCAAGCGGAAAGCGACAGNA 1815  
1816 GAATCTGAGCTGAGAGATCTCGGCTTGTCTGCTGAAGAAATCAAGCGGAAAGCGACAGAA 1757  
1816 TGTTCAGATAAATCTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAAGTGGGAGAAATATCG 1875  
1756 TGTTCAGATAAATCTGTTGAAGGCTAAGGATCAAGGGAAGCCTGAAGTGGGAGAAATATCG 1697  
1876 AAATCTGGAAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAGCTCCAGAGAAATCTG 1935  
1696 AAATCTGGAAGATCAATGCTGAGCAGAGCTCAAAATTCAGGAGCTCCAGAGAAATCTG 1637  
1936 GAGAGGCT-----GCA 1947  
1636 GAGAGGCTGTTAAAGCCAGCAGGAGGCCACCGAGCTGCTGCAGNAATATCCGCCAGGCA 1577  
1948 AAGGAGCGAGCCGAGAGGAGTGGAGAACTGCAGAACCGAGAGGATCTTTCTGAAGGC 2007  
1576 AAGGAGCGAGCCGAGAGGAGTGGAGAACTGCAGAACCGAGAGGATCTTTCTGAAGGC 1517  
2008 ATCAGAAAGAGCTGTTGAGCTGAGGAGCGGAGCTTCTCTGAGAGCAAGGTAAAG 2067  
1516 ATCAGAAAGAGCTGTTGAGCTGAGGAGCGGAGCTTCTCTGAGAGCAAGGTAAAG 1457  
2068 AGACTAGAGACCATGGAGCTTAGAGAAACAGACTGAAGGATGACATCCAGAGCAAAATCC 2127  
1456 AGACTAGAGACCATGGAGCTTAGAGAAACAGACTGAAGGATGACATCCAGAGCAAAATCC 1397  
2128 CAACAGATCAGAGATGGCTGATAAATTTCTGGAGCTGAGAGAAACATCGGAGGCC 2187  
1396 CAACAGATCAGAGATGGCTGATAAATTTCTGGAGCTGAGAGAAACATCGGAGGCC 1337  
2188 CAAGTCTCAGCCAGCACCTAGAAAGTGCACCTGAAACAGAGAGCAGCACTATGA -GGA 2246  
1336 CAAGTCTCAGCCAGCACCTAGAAAGTGCACCTGAAACAGAGAGCAGCACTATGAGGA 1277  
2247 AAAGATTAAAGTGTGGACAAATCAGATAAAGAAACCTGGGTGACAGAGAGCACTGGA 2306  
1276 AAAGATTAAAGTGTGGACAAATCAGATAAAGAAACCTGGGTGACAGAGAGCACTGGA 1217  
2307 GAACATGATCGAGACACAGAGGAGGAGCCCATGAGAGGCGCAAAATTTCTCAGCGAACA 2366  
1216 GAACATGATCGAGACACAGAGGAGGAGCCCATGAGAGGCGCAAAATTTCTCAGCGAACA 1157  
2367 GAAGGGGATGATCAATGCTATGGAATTCAGAGTACAGATCCCTGGAACAGAGGATGTGA 2426  
1156 GAAGGGGATGATCAATGCTATGGAATTCAGAGTACAGATCCCTGGAACAGAGGATGTGA 1097  
2427 ACTGTCTGAGCCAAATAAATTCAGCAATAGCAGTCTTTTACCAGAGGACATGAA 2486  
1096 ACTGTCTGAGCCAAATAAATTCAGCAATAGCAGTCTTTTACCAGAGGACATGAA 1037  
2487 GGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTACCTGGAGACACAGGC 2546  
1036 GGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTACCTGGAGACACAGGC 977  
2547 TGGGAATTTGGAGGCCAGAACCGGAAACTGAGGAGGAGCTGGAGAGATCAGCCACCA 2606  
976 TGGGAATTTGGAGGCCAGAACCGGAAACTGAGGAGGAGCTGGAGAGATCAGCCACCA 917  
2607 AGACCAGTGAAGAAATCGGCTGCTGGAATCTGGAGCAAGATTCGGGAGGTCAGTCT 2666  
916 AGACCAGTGAAGAAATCGGCTGCTGGAATCTGGAGCAAGATTCGGGAGGTCAGTCT 857  
2667 AGACACGAGGAGCAAAATCGAGTCAAGAGCCAGCTCAGAGCTCAGCTCTCCCT 2726

856 AGACACGAGGAGCAGAAACTGGAGCTCAAGCGCCAGCTCACAGAGCTACAGCTCTCCCT 797  
2727 GCAGGAGCGGAGTCAAGTTGACAGCCCTGCAGGCTGCACGGGCGCCCTGGAGAGCCA 2786  
796 GCAGGAGCGGAGTCAAGTTGACAGCCCTGCAGGCTGCACGGGCGCCCTGGAGAGCCA 737  
2787 GCTTCGCCAGCGAAGACAGAGCTGGAAGAGACACACAGCAGAAAGCTGAAAGAGAGATCCA 2846  
736 GCTTCGCCAGCGAAGACAGAGCTGGAAGAGACACACAGCAGAAAGCGAAGAGAGATCCA 677  
2847 GGCACTCACGCGACATAGAGATGAATTCAGCGCAAAATTTGATGCTCTTCTGAACAGCTG 2906  
676 GGCACTCACGCGACATAGAGATGAATTCAGCGCAAAATTTGATGCTCTTCTGAACAGCTG 617  
2907 TACTGTATCACAGACCTGGAGGAGCAGCTAAACAGCTCAGCAGGAGCAACACGCTGAAT 2966  
616 TACTGTATCACAGACCTGGAGGAGCAGCTAAACAGCTCAGCAGGAGCAACACGCTGAAT 557  
2967 CAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCCAAACGACGA 3026  
556 CAACAAACCAAACTTCTACTTGTCCAAACAACTCGATGAGGCTTCTGGCGCCAAACGACGA 497  
3027 GATTGTAACTCGAAGTGAAGTGAACCATCTCCGCGCGGAGATCACGGAAACGAGAGAT 3086  
496 GATTGTAACTCGAAGTGAAGTGAACCATCTCCGCGCGGAGATCACGGAAACGAGAGAT 437  
3087 GCAGCTTACCAGCCAGAGCAAAACGATGAGGCTTCTGAAGACCAACGTCACCATGTGGA 3146  
436 GCAGCTTACCAGCCAGAGCAAAACGATGAGGCTTCTGAAGACCAACGTCACCATGTGGA 377  
3147 GGAACAGGCTATGGAATTTGAGGCGCTTAAACGATGAGCTCTAGAGAAAGAGGCGAGTG 3206  
376 GGAACAGGCTATGGAATTTGAGGCGCTTAAACGATGAGCTCTAGAGAAAGAGGCGAGTG 317  
3207 GGAGGCTTGAGGAGGCTCTCTGGGTGATGAGAAATCCAGTCTTGGAGTGTCCGCTTCGAGA 3266  
316 GGAGGCTTGAGGAGGCTCTCTGGGTGATGAGAAATCCAGTCTTGGAGTGTCCGCTTCGAGA 257  
3267 GCTCCAGAGAAATGCTGACACCCAGAGAAACAGAGAGGCGGAGAGCCGATCAGCGGATCAC 3326  
256 GCTCCAGAGAAATGCTGACACCCAGAGAAACAGAGAGGCGGAGAGCCGATCAGCGGATCAC 197  
3327 CGAGTCTCGCAGGCTGAGGCTGAGGCTGAGGAGCAAGGCTGAGATTTCTCGCTCT 3386  
196 CGAGTCTCGCAGGCTGAGGCTGAGGCTGAGGAGCAAGGCTGAGATTTCTCGCTCT 137  
3387 GCAGCAGGCTCTCAAGAGCAGAGCTGGAAGGCGGAGAGCTCTCTGACAAAGCTCAATGA 3446  
136 GCAGCAGGCTCTCAAGAGCAGAGCTGGAAGGCGGAGAGCTCTCTGACAAAGCTCAATGA 77  
3447 CCTGGAGAGAGCATGCTATGTTGAAATGAATGCCCGAGGCTTACAGCAGAGCTGGA 3506  
76 CCTGGAGA-----AGAGCTGGA 59  
3507 GACTGAAACGAGAGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAATTTACAGCAGCAG 3564  
58 GACTGAAACGAGAGCTCAACAGAGGCTTCTGGAGAGCAAGCCAAATTTACAGCAGCAG 1

## RESULT 15

ABA08361

ID ABA08361 standard; cDNA; 3131 BP.

XX ABA08361;

XX 11-JAN-2002 (first entry)

XX Human RHO/RAC effector homologue-encoding cDNA, SEQ ID NO:137.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;

XX haematopoiesis regulation; tissue growth; immunomodulator; activin;

XX inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;

proliferation; metastasis; cancer; tumour; haematopoietic disorder; myeloid cell disorder; lymphoid cell disorder; asthma; arthritis; chronic inflammatory condition; proliferative retinopathy; atherosclerosis; coronary heart disease; arterial ischaemia; bone disorder; osteoporosis; vascular growth disorder; tissue regeneration; wound healing; infection; immune disorder; cell culture; drug screening; gene therapy; antiinflammatory; antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic; cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial; antifungal; vulnery; antilulcer; ss.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US003800.

03-FEB-2000; 2000US-00496914.

27-FEB-2000; 2000US-00560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT; WPI; 2001-457740/49.

P-PSDB; ABB11117.

Human proteins and DNA encoding sequences useful for preventing, treating or ameliorating a medical condition in a mammalian subject e.g. arthritis and cancer.

Claim 1; Page 387-388; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention

SQ Sequence 3131 BP; 861 A; 835 C; 806 G; 629 T; 0 U; 0 Other;

Query March 41.63; Score 2564.8; DB 4; Length 3131;  
Best Local Similarity 93.44; Pred. No. 0;  
Matches 2781; Conservative 0; Mismatches 2; Indels 195; Gaps

QY 2857 GCACATAGATGAATCAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAACT 2916  
DB 1 GCACATAGATGAATCCAGCGCAAAATTTGATGCTCTTCGTAAACAGCTGTACTGTAACT 60

QY 2917 ACAGACCTGGAGGAGCAGCTAAACCCAGCTGACCGAGGACAAACGCTGAACCTCAACAAACAA 2976  
DB 61 ACAGACCTGGAGGAGCAGCTAAACCCAGCTGACCGAGGACAAACGCTGAACCTCAACAAACAA 120

QY 2977 AACCTTCTACTTTGTCCAAACCAACTCGATGAGGCTTCTCGCGCCAAACGACGAGATTGTACAA 3036  
DB 121 AACCTTCTACTTTGTCCAAACCAACTCGATGAGGCTTCTCGCGCCAAACGACGAGATTGTACAA 180

QY 3037 CTGCGAAGTGAAGTGGACCATCTCCCGCCGGGAGATACCGGAAACGAGAGATGCAGCTTACC 3096  
DB 181 CTGCGAAGTGAAGTGGACCATCTCCCGCCGGGAGATACCGGAAACGAGAGATGCAGCTTACC 240

QY 3097 AGCCAGAGCAACGATGGAGGCTCTGAAGACCAACGTCACCATGCTGGAGGAAACAGGTC 3156  
DB 241 AGCCAGAGCAACGATGGAGGCTCTGAAGACCAACGTCACCATGCTGGAGGAAACAGGTC 300

QY 3157 ATGGATTGGAGGCCCTAAACGATGAGCTGCTAGAAAAAGAGCGGCGAGTGGAGGCCCTGG 3216  
DB 301 ATGGATTGGAGGCCCTAAACGATGAGCTGCTAGAAAAAGAGCGGCGAGTGGAGGCCCTGG 360

QY 3217 AGGAGCGTCTGGGTGATGAGAAATCCAGTTTGAAGTGTGGGTTCGAGAGCTGCAGAGA 3276  
DB 361 AGGAGCGTCTGGGTGATGAGAAATCCAGTTTGAAGTGTGGGTTCGAGAGCTGCAGAGA 420

QY 3277 ATGCTGACACCCGAGAAACAGAGCAGCGGCGAGAGCCGATCAGCGGATCAACGAGTCTCGC 3336  
DB 421 ATGCTGACACCCGAGAAACAGAGCAGCGGCGAGAGCCGATCAGCGGATCAACGAGTCTCGC 480

QY 3337 CAGTGTGGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTTCGCTTCGCTGCAGCAGGCT 3396  
DB 481 CAGTGTGGAGCTGGCAGTGAAGGAGCAACAGGCTGAGATTTCGCTTCGCTGCAGCAGGCT 540

QY 3397 CTCAAAGAGCAGAGAGCTGAAGCGCGAGAGGCTCTCTCACAAGCTCAATGACTCGGAGAAG 3456  
DB 541 CTCAAAGAGCAGAGAGCTGAAGCGCGAGAGGCTCTCTCACAAGCTCAATGACTCGGAGAAG 600

QY 3457 AAGCATGCTATGCTTGAATGAATGCCGAGGCTTACAGCAGAGAGCTGGAGACTGAACGA 3516  
DB 601 AAGCATGCTATGCTTGAATGAATGCCGAGGCTTACAGCAGAGAGCTGGAGACTGAACGA 660

QY 3517 GAGCTCAACAGAGGCTCTGGAAGAGCAAGCCAAATACAGCAGCAGATGGAGCTGCAG 3576  
DB 661 GAGCTCAACAGAGGCTCTGGAAGAGCAAGCCAAATACAGCAGCAGATGGAGCTGCAG 720

QY 3577 AAAAATCACATTTTCCGCTCTGACTCAAGGAGCTGCAAGAAGCTCTAGATCGGGCTGATCTA 3636  
DB 721 AAAAATCACATTTTCCGCTCTGACTCAAGGAGCTGCAAGAAGCTCTAGATCGGGCTGATCTA 780

QY 3637 CTGAAGACAGAAAGAGTACTTGGAGTATCAGCTGGAACCAATTCAGGTTCTCTATCT 3696  
DB 781 CTGAAGACAGAAAGAGTACTTGGAGTATCAGCTGGAACCAATTCAGGTTCTCTATCT 840

QY 3697 CATGAAAGGTGAAATGGAAGGCACTATTCTCAACAAACCAAACTCATTTGATTTCTG 3756  
DB 841 CATGAAAGGTGAAATGGAAGGCACTATTCTCAACAAACCAAACTCATTTGATTTCTG 900

QY 3757 CAAGCCAAAATGGACCAACTGCTTAAAGAAAGAGGGTTTATTAGTCGACGGAAGAG 3816  
DB 901 CAAGCCAAAATGGACCAACTGCTTAAAGAAAGAGGGTTTATTAGTCGACGGAAGAG 934

QY 3817 GACCTGCTTTACCCACACAGGTTCTCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAG 3876  
DB 935 -----AGGTTCTCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAG 975

